

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 12:31:50 ; Search time 110 Seconds  
(without alignments)

889.156 Million cell updates/sec

Title: US-10-770-117-4

Perfect score: 1040

Sequence: 1 MSTNPKPKQKKNTRRPO.....CSPSIFLLALLSLCTVPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040	100.0	3011	1	P26664 h genome po
2	1031	99.1	191	2	Q68113 hepatitis c
3	1031	99.1	191	2	Q68123 hepatitis c
4	1031	99.1	191	2	Q68124 hepatitis c
5	1031	99.1	191	2	Q68139 hepatitis c
6	1031	99.1	191	2	Q68149 hepatitis c
7	1031	99.1	191	2	Q68149 hepatitis c
8	1031	99.1	191	2	Q68149 hepatitis c
9	1031	99.1	191	2	Q68149 hepatitis c
10	1031	99.1	191	2	Q68149 hepatitis c
11	1031	99.1	191	2	Q68149 hepatitis c
12	1031	99.1	191	2	Q68149 hepatitis c
13	1031	99.1	191	2	Q68149 hepatitis c
14	1031	99.1	191	2	Q68149 hepatitis c
15	1031	99.1	191	2	Q68149 hepatitis c
16	1031	99.1	191	2	Q68149 hepatitis c
17	1031	99.1	191	2	Q68149 hepatitis c
18	1031	99.1	191	2	Q68149 hepatitis c
19	1031	99.1	191	2	Q68149 hepatitis c
20	1031	99.1	191	2	Q68149 hepatitis c
21	1031	99.1	191	2	Q68149 hepatitis c
22	1031	99.1	191	2	Q68149 hepatitis c
23	1031	99.1	191	2	Q68149 hepatitis c
24	1031	99.1	191	2	Q68149 hepatitis c
25	1031	99.1	191	2	Q68149 hepatitis c
26	1031	99.1	191	2	Q68149 hepatitis c
27	1031	99.1	191	2	Q68149 hepatitis c
28	1031	99.1	191	2	Q68149 hepatitis c
29	1031	99.1	191	2	Q68149 hepatitis c
30	1031	99.1	191	2	Q68149 hepatitis c
31	1031	99.1	191	2	Q68149 hepatitis c

32 1031 99.1 191 2 Q6QYF6  
33 1031 99.1 191 2 Q6QYF7  
34 1031 99.1 191 2 Q6QYF9  
35 1031 99.1 191 2 Q6QYGO  
36 1031 99.1 191 2 Q6QYGS  
37 1031 99.1 191 2 Q6QYJ7  
38 1031 99.1 191 2 Q6QYJ9  
39 1031 99.1 191 2 Q6QYK0  
40 1031 99.1 191 2 Q6QYK2  
41 1031 99.1 191 2 Q6QYK3  
42 1031 99.1 191 2 Q6QYK6  
43 1031 99.1 191 2 Q6QYK7  
44 1031 99.1 191 2 Q6QYK8  
45 1031 99.1 191 2 Q6QYK9

#### ALIGNMENTS

RESULT 1  
ID POLG HCV1 STANDARD; PRT; 3011 AA.  
AC P26664;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Genome polyprotein [Contains: Capsid protein C (Core protein) (P22); Envelope glycoprotein E1 (GP32) (GP35); Envelope glycoprotein E2 (GP68) (GP70) (NS1); Protein P7; Nonstructural protein NS2 (P21) (P22) (P23) (P24) (P25) (P26) (P27) (P28) (P29) (P30) (P31) (P32) (P33) (P34) (P35) (P36) (P37) (P38) (P39) (P40) (P41) (P42) (P43) (P44) (P45) (P46) (P47) (P48) (P49) (P50) (P51) (P52) (P53) (P54) (P55) (P56) (P57) (P58) (P59) (P60) (P61) (P62) (P63) (P64) (P65) (P66) (P67) (P68) (P69) (P70) (P71) (P72) (P73) (P74) (P75) (P76) (P77) (P78) (P79) (P80) (P81) (P82) (P83) (P84) (P85) (P86) (P87) (P88) (P89) (P90) (P91) (P92) (P93) (P94) (P95) (P96) (P97) (P98) (P99) (P100) (P101) (P102) (P103) (P104) (P105) (P106) (P107) (P108) (P109) (P110) (P111) (P112) (P113) (P114) (P115) (P116) (P117) (P118) (P119) (P120) (P121) (P122) (P123) (P124) (P125) (P126) (P127) (P128) (P129) (P130) (P131) (P132) (P133) (P134) (P135) (P136) (P137) (P138) (P139) (P140) (P141) (P142) (P143) (P144) (P145) (P146) (P147) (P148) (P149) (P150) (P151) (P152) (P153) (P154) (P155) (P156) (P157) (P158) (P159) (P160) (P161) (P162) (P163) (P164) (P165) (P166) (P167) (P168) (P169) (P170) (P171) (P172) (P173) (P174) (P175) (P176) (P177) (P178) (P179) (P180) (P181) (P182) (P183) (P184) (P185) (P186) (P187) (P188) (P189) (P190) (P191) (P192) (P193) (P194) (P195) (P196) (P197) (P198) (P199) (P200) (P201) (P202) (P203) (P204) (P205) (P206) (P207) (P208) (P209) (P210) (P211) (P212) (P213) (P214) (P215) (P216) (P217) (P218) (P219) (P220) (P221) (P222) (P223) (P224) (P225) (P226) (P227) (P228) (P229) (P230) (P231) (P232) (P233) (P234) (P235) (P236) (P237) (P238) (P239) (P240) (P241) (P242) (P243) (P244) (P245) (P246) (P247) (P248) (P249) (P250) (P251) (P252) (P253) (P254) (P255) (P256) (P257) (P258) (P259) (P260) (P261) (P262) (P263) (P264) (P265) (P266) (P267) (P268) (P269) (P270) (P271) (P272) (P273) (P274) (P275) (P276) (P277) (P278) (P279) (P280) (P281) (P282) (P283) (P284) (P285) (P286) (P287) (P288) (P289) (P290) (P291) (P292) (P293) (P294) (P295) (P296) (P297) (P298) (P299) (P300) (P301) (P302) (P303) (P304) (P305) (P306) (P307) (P308) (P309) (P310) (P311) (P312) (P313) (P314) (P315) (P316) (P317) (P318) (P319) (P320) (P321) (P322) (P323) (P324) (P325) (P326) (P327) (P328) (P329) (P330) (P331) (P332) (P333) (P334) (P335) (P336) (P337) (P338) (P339) (P340) (P341) (P342) (P343) (P344) (P345) (P346) (P347) (P348) (P349) (P350) (P351) (P352) (P353) (P354) (P355) (P356) (P357) (P358) (P359) (P360) (P361) (P362) (P363) (P364) (P365) (P366) (P367) (P368) (P369) (P370) (P371) (P372) (P373) (P374) (P375) (P376) (P377) (P378) (P379) (P380) (P381) (P382) (P383) (P384) (P385) (P386) (P387) (P388) (P389) (P390) (P391) (P392) (P393) (P394) (P395) (P396) (P397) (P398) (P399) (P400) (P401) (P402) (P403) (P404) (P405) (P406) (P407) (P408) (P409) (P410) (P411) (P412) (P413) (P414) (P415) (P416) (P417) (P418) (P419) (P420) (P421) (P422) (P423) (P424) (P425) (P426) (P427) (P428) (P429) (P430) (P431) (P432) (P433) (P434) (P435) (P436) (P437) (P438) (P439) (P440) (P441) (P442) (P443) (P444) (P445) (P446) (P447) (P448) (P449) (P450) (P451) (P452) (P453) (P454) (P455) (P456) (P457) (P458) (P459) (P460) (P461) (P462) (P463) (P464) (P465) (P466) (P467) (P468) (P469) (P470) (P471) (P472) (P473) (P474) (P475) (P476) (P477) (P478) (P479) (P480) (P481) (P482) (P483) (P484) (P485) (P486) (P487) (P488) (P489) (P490) (P491) (P492) (P493) (P494) (P495) (P496) (P497) (P498) (P499) (P500) (P501) (P502) (P503) (P504) (P505) (P506) (P507) (P508) (P509) (P510) (P511) (P512) (P513) (P514) (P515) (P516) (P517) (P518) (P519) (P520) (P521) (P522) (P523) (P524) (P525) (P526) (P527) (P528) (P529) (P530) (P531) (P532) (P533) (P534) (P535) (P536) (P537) (P538) (P539) (P540) (P541) (P542) (P543) (P544) (P545) (P546) (P547) (P548) (P549) (P550) (P551) (P552) (P553) (P554) (P555) (P556) (P557) (P558) (P559) (P560) (P561) (P562) (P563) (P564) (P565) (P566) (P567) (P568) (P569) (P570) (P571) (P572) (P573) (P574) (P575) (P576) (P577) (P578) (P579) (P580) (P581) (P582) (P583) (P584) (P585) (P586) (P587) (P588) (P589) (P590) (P591) (P592) (P593) (P594) (P595) (P596) (P597) (P598) (P599) (P600) (P601) (P602) (P603) (P604) (P605) (P606) (P607) (P608) (P609) (P610) (P611) (P612) (P613) (P614) (P615) (P616) (P617) (P618) (P619) (P620) (P621) (P622) (P623) (P624) (P625) (P626) (P627) (P628) (P629) (P630) (P631) (P632) (P633) (P634) (P635) (P636) (P637) (P638) (P639) (P640) (P641) (P642) (P643) (P644) (P645) (P646) (P647) (P648) (P649) (P650) (P651) (P652) (P653) (P654) (P655) (P656) (P657) (P658) (P659) (P660) (P661) (P662) (P663) (P664) (P665) (P666) (P667) (P668) (P669) (P670) (P671) (P672) (P673) (P674) (P675) (P676) (P677) (P678) (P679) (P680) (P681) (P682) (P683) (P684) (P685) (P686) (P687) (P688) (P689) (P690) (P691) (P692) (P693) (P694) (P695) (P696) (P697) (P698) (P699) (P700) (P701) (P702) (P703) (P704) (P705) (P706) (P707) (P708) (P709) (P710) (P711) (P712) (P713) (P714) (P715) (P716) (P717) (P718) (P719) (P720) (P721) (P722) (P723) (P724) (P725) (P726) (P727) (P728) (P729) (P730) (P731) (P732) (P733) (P734) (P735) (P736) (P737) (P738) (P739) (P740) (P741) (P742) (P743) (P744) (P745) (P746) (P747) (P748) (P749) (P750) (P751) (P752) (P753) (P754) (P755) (P756) (P757) (P758) (P759) (P760) (P761) (P762) (P763) (P764) (P765) (P766) (P767) (P768) (P769) (P770) (P771) (P772) (P773) (P774) (P775) (P776) (P777) (P778) (P779) (P780) (P781) (P782) (P783) (P784) (P785) (P786) (P787) (P788) (P789) (P790) (P791) (P792) (P793) (P794) (P795) (P796) (P797) (P798) (P799) (P800) (P801) (P802) (P803) (P804) (P805) (P806) (P807) (P808) (P809) (P810) (P811) (P812) (P813) (P814) (P815) (P816) (P817) (P818) (P819) (P820) (P821) (P822) (P823) (P824) (P825) (P826) (P827) (P828) (P829) (P830) (P831) (P832) (P833) (P834) (P835) (P836) (P837) (P838) (P839) (P840) (P841) (P842) (P843) (P844) (P845) (P846) (P847) (P848) (P849) (P850) (P851) (P852) (P853) (P854) (P855) (P856) (P857) (P858) (P859) (P860) (P861) (P862) (P863) (P864) (P865) (P866) (P867) (P868) (P869) (P870) (P871) (P872) (P873) (P874) (P875) (P876) (P877) (P878) (P879) (P880) (P881) (P882) (P883) (P884) (P885) (P886) (P887) (P888) (P889) (P890) (P891) (P892) (P893) (P894) (P895) (P896) (P897) (P898) (P899) (P900) (P901) (P902) (P903) (P904) (P905) (P906) (P907) (P908) (P909) (P910) (P911) (P912) (P913) (P914) (P915) (P916) (P917) (P918) (P919) (P920) (P921) (P922) (P923) (P924) (P925) (P926) (P927) (P928) (P929) (P930) (P931) (P932) (P933) (P934) (P935) (P936) (P937) (P938) (P939) (P940) (P941) (P942) (P943) (P944) (P945) (P946) (P947) (P948) (P949) (P950) (P951) (P952) (P953) (P954) (P955) (P956) (P957) (P958) (P959) (P960) (P961) (P962) (P963) (P964) (P965) (P966) (P967) (P968) (P969) (P970) (P971) (P972) (P973) (P974) (P975) (P976) (P977) (P978) (P979) (P980) (P981) (P982) (P983) (P984) (P985) (P986) (P987) (P988) (P989) (P990) (P991) (P992) (P993) (P994) (P995) (P996) (P997) (P998) (P999) (P1000) (P1001) (P1002) (P1003) (P1004) (P1005) (P1006) (P1007) (P1008) (P1009) (P1010) (P1011) (P1012) (P1013) (P1014) (P1015) (P1016) (P1017) (P1018) (P1019) (P1020) (P1021) (P1022) (P1023) (P1024) (P1025) (P1026) (P1027) (P1028) (P1029) (P1030) (P1031) (P1032) (P1033) (P1034) (P1035) (P1036) (P1037) (P1038) (P1039) (P1040) (P1041) (P1042) (P1043) (P1044) (P1045) (P1046) (P1047) (P1048) (P1049) (P1050) (P1051) (P1052) (P1053) (P1054) (P1055) (P1056) (P1057) (P1058) (P1059) (P1060) (P1061) (P1062) (P1063) (P1064) (P1065) (P1066) (P1067) (P1068) (P1069) (P1070) (P1071) (P1072) (P1073) (P1074) (P1075) (P1076) (P1077) (P1078) (P1079) (P1080) (P1081) (P1082) (P1083) (P1084) (P1085) (P1086) (P1087) (P1088) (P1089) (P1090) (P1091) (P1092) (P1093) (P1094) (P1095) (P1096) (P1097) (P1098) (P1099) (P1100) (P1101) (P1102) (P1103) (P1104) (P1105) (P1106) (P1107) (P1108) (P1109) (P1110) (P1111) (P1112) (P1113) (P1114) (P1115) (P1116) (P1117) (P1118) (P1119) (P1120) (P1121) (P1122) (P1123) (P1124) (P1125) (P1126) (P1127) (P1128) (P1129) (P1130) (P1131) (P1132) (P1133) (P1134) (P1135) (P1136) (P1137) (P1138) (P1139) (P1140) (P1141) (P1142) (P1143) (P1144) (P1145) (P1146) (P1147) (P1148) (P1149) (P1150) (P1151) (P1152) (P1153) (P1154) (P1155) (P1156) (P1157) (P1158) (P1159) (P1160) (P1161) (P1162) (P1163) (P1164) (P1165) (P1166) (P1167) (P1168) (P1169) (P1170) (P1171) (P1172) (P1173) (P1174) (P1175) (P1176) (P1177) (P1178) (P1179) (P1180) (P1181) (P1182) (P1183) (P1184) (P1185) (P1186) (P1187) (P1188) (P1189) (P1190) (P1191) (P1192) (P1193) (P1194) (P1195) (P1196) (P1197) (P1198) (P1199) (P1200) (P1201) (P1202) (P1203) (P1204) (P1205) (P1206) (P1207) (P1208) (P1209) (P1210) (P1211) (P1212) (P1213) (P1214) (P1215) (P1216) (P1217) (P1218) (P1219) (P1220) (P1221) (P1222) (P1223) (P1224) (P1225) (P1226) (P1227) (P1228) (P1229) (P1230) (P1231) (P1232) (P1233) (P1234) (P1235) (P1236) (P1237) (P1238) (P1239) (P1240) (P1241) (P1242) (P1243) (P1244) (P1245) (P1246) (P1247) (P1248) (P1249) (P1250) (P1251) (P1252) (P1253) (P1254) (P1255) (P1256) (P1257) (P1258) (P1259) (P1260) (P1261) (P1262) (P1263) (P1264) (P1265) (P1266) (P1267) (P1268) (P1269) (P1270) (P1271) (P1272) (P1273) (P1274) (P1275) (P1276) (P1277) (P1278) (P1279) (P1280) (P1281) (P1282) (P1283) (P1284) (P1285) (P1286) (P1287) (P1288) (P1289) (P1290) (P1291) (P1292) (P1293) (P1294) (P1295) (P1296) (P1297) (P1298) (P1299) (P1300) (P1301) (P1302) (P1303) (P1304) (P1305) (P1306) (P1307) (P1308) (P1309) (P1310) (P1311) (P1312) (P1313) (P1314) (P1315) (P1316) (P1317) (P1318) (P1319) (P1320) (P1321) (P1322) (P1323) (P1324) (P1325) (P1326) (P1327) (P1328) (P1329) (P1330) (P1331) (P1332) (P1333) (P1334) (P1335) (P1336) (P1337) (P1338) (P1339) (P1340) (P1341) (P1342) (P1343) (P1344) (P1345) (P1346) (P1347) (P1348) (P1349) (P1350) (P1351) (P1352) (P1353) (P1354) (P1355) (P1356) (P1357) (P1358) (P1359) (P1360) (P1361) (P1362) (P1363) (P1364) (P1365) (P1366) (P1367) (P1368) (P1369) (P1370) (P1371) (P1372) (P1373) (P1374) (P1375) (P1376) (P1377) (P1378) (P1379) (P1380) (P1381) (P1382) (P1383) (P1384) (P1385) (P1386) (P1387) (P1388) (P1389) (P1390) (P1391) (P1392) (P1393) (P1394) (P1395) (P1396) (P1397) (P1398) (P1399) (P1400) (P1401) (P1402) (P1403) (P1404) (P1405) (P1406) (P1407) (P14

[illegible]

[illegible]





```

RP SEQUENCE FROM N.A.
RX MEDLINE=93376778; PubMed=8396266;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "At least 12 genotypes of hepatitis C virus predicted by sequence
RL analysis of the putative E1 gene of isolates collected worldwide.";
RR Proc. Natl. Acad. Sci. U.S.A. 90:8234-8238(1993).
DR EMBL; U10232; AAA21071.1; -.
DR PIR; P00804; P00804.
DR HSSP; O8JYS1; ICWX.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
KW Polyprotein.
FT CHAIN 1 >191 core protein.
FT NON TER 191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 99.1%; Score 1031; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2e-79;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKPKQKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
Db 1 MSTNPKPKQKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60

Qy 61 RRQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGSPPSWGPTDPRRRSRNLG 120
Db 61 RRQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGSPPSWGPTDPRRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSLCTVPASA 191
Db 181 LLSLCTVPASA 191

RESULT 7
Q6QY91 ID Q6QY91 PRELIMINARY; PRT; 191 AA.
AC Q6QY91;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522326; AAS15488.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
FT NON TER 191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 99.1%; Score 1031; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2e-79;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKPKQKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
Db 1 MSTNPKPKQKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60

Qy 61 RRQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGSPPSWGPTDPRRRSRNLG 120
Db 61 RRQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGSPPSWGPTDPRRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSLCTVPASA 191
Db 181 LLSLCTVPASA 191

RESULT 9
Q6QY99 ID Q6QY99 PRELIMINARY; PRT; 191 AA.
AC Q6QY99;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522326; AAS15488.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
FT NON TER 191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 99.1%; Score 1031; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2e-79;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKPKQKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
Db 1 MSTNPKPKQKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60

Qy 61 RRQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGSPPSWGPTDPRRRSRNLG 120
Db 61 RRQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGSPPSWGPTDPRRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSLCTVPASA 191
Db 181 LLSLCTVPASA 191

```

```
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522318; AAS15480.1; -
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON_TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 99.1%; Score 1031; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2e-79;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNTRRRPQDVKFGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 MSTNPKPQRTKRNTRRRPQDVKFGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60

QY 61 RRQIPKARRPEGRTWAOQPGYPWPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
61 RRQIPKARRPEGRTWAOQPGYPWPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191
Db |||||:|||||
181 LLSCLTVPASA 191

RESULT 10
Q6QYAL PRELIMINARY; PRT; 191 AA.
ID Q6QYAL
AC Q6QYAL;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522315; AAS15477.1; -
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON_TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 99.1%; Score 1031; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2e-79;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNTRRRPQDVKFGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 MSTNPKPQRTKRNTRRRPQDVKFGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60

QY 61 RRQIPKARRPEGRTWAOQPGYPWPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
61 RRQIPKARRPEGRTWAOQPGYPWPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191
Db |||||:|||||
181 LLSCLTVPASA 191

RESULT 11
Q6QYAL PRELIMINARY; PRT; 191 AA.
ID Q6QYAL
AC Q6QYAL;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522315; AAS15477.1; -
DR GO; GO:0019028; C:Viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON_TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 99.1%; Score 1031; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2e-79;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNTRRRPQDVKFGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1 MSTNPKPQRTKRNTRRRPQDVKFGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60

QY 61 RRQIPKARRPEGRTWAOQPGYPWPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
61 RRQIPKARRPEGRTWAOQPGYPWPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191
Db |||||:|||||
181 LLSCLTVPASA 191

RESULT 12
Q6QYAL PRELIMINARY; PRT; 191 AA.
ID Q6QYAL
AC Q6QYAL;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AY522311; AAS15473.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 99.1%; Score 1031; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2e-79;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNRTNRRPDQVKGPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKPQKTKNTNRRPDQVKGPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60

Qy 61 RRQPIKARRPEGRGTWAQPGYPWPLYNGEGCGWAGWLLSPRGSPSWGPTDPRRSRNLG 120
Db 61 RRQPIKARRPEGRGTWAQPGYPWPLYNGEGCGWAGWLLSPRGSPSWGPTDPRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDCGVNATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDCGVNATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 13
Q6QYA7 : PRELIMINARY; PRT; 191 AA.
ID Q6QYA7 : PRELIMINARY; PRT; 191 AA.
AC Q6QYA7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E.; Qin H.; Shire N.; Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522310; AAS15472.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 99.1%; Score 1031; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2e-79;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNRTNRRPDQVKGPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKPQKTKNTNRRPDQVKGPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60

Qy 61 RRQPIKARRPEGRGTWAQPGYPWPLYNGEGCGWAGWLLSPRGSPSWGPTDPRRSRNLG 120
Db 61 RRQPIKARRPEGRGTWAQPGYPWPLYNGEGCGWAGWLLSPRGSPSWGPTDPRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDCGVNATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDCGVNATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 14
Q6QYB1 : PRELIMINARY; PRT; 191 AA.
ID Q6QYB1 : PRELIMINARY; PRT; 191 AA.
AC Q6QYB1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E.; Qin H.; Shire N.; Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522306; AAS15468.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 99.1%; Score 1031; DB 2; Length 191;
Best Local Similarity 99.0%; Pred. No. 2e-79;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNRTNRRPDQVKGPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKPQKTKNTNRRPDQVKGPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60

Qy 61 RRQPIKARRPEGRGTWAQPGYPWPLYNGEGCGWAGWLLSPRGSPSWGPTDPRRSRNLG 120
Db 61 RRQPIKARRPEGRGTWAQPGYPWPLYNGEGCGWAGWLLSPRGSPSWGPTDPRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDCGVNATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDCGVNATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 15
Q6QYB2 : PRELIMINARY; PRT; 191 AA.
ID Q6QYB2 : PRELIMINARY; PRT; 191 AA.
AC Q6QYB2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E.; Qin H.; Shire N.; Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522305; AAS15467.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
```

DR Pfam: PF01543; HCV capsid; 1.  
DR Pfam: PF01542; HCV\_core; 1.  
FT NON\_TER 191-191  
SQ SEQUENCE 191 AA; 20766 MW; 4506596P21B4A893 CRC64;

Query Match 99.1%; Score 1031; DB 2; Length 191;  
Best Local Similarity 99.0%; Pred. No. 2e-79;  
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MSTNPKQKKNKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG	60
Db	1	MSTNPKPQRTKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG	60
Qy	61	RRQPIPKARRPEGRWAOQGYWPLYGNECCGAGWLLSPRGSRPSPWGPTDPRRRSRNLG	120
Db	61	RRQPIPKARRPEGRWAOQGYWPLYGNECCGAGWLLSPRGSRPSPWGPTDPRRRSRNLG	120
Qy	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA	180
Db	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA	180
Qy	181	LLSCLTVPASA	191
Db	181	LLSCLTVPASA	191

Search completed: August 19, 2005, 12:43:13  
Job time : 111 secs



Db	1	MSTNPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Qy	61	RRQIPKARRPEGRTWAOQGYWPWLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
Db	61	RRQIPKARRPEGRTWAOQGYWPWLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
Qy	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNYYATGNLPGCSFSIFLLA 180
Db	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNYYATGNLPGCSFSIFLLA 180
Qy	181	LLSCLTVPASA 191
Db	181	LLSCLTVPASA 191
RESULT 2		
GNWVCJ		
genome polyprotein - hepatitis C virus (strain H)		
N;Contains: capsid protein C; envelope protein M; hepatitis C virus (strain H)		
Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5		
C;Species: hepatitis C virus		
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004		
A;Accession: A36814; A41546		
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.		
submitted to GenBank, July 1992		
A;Description: Genomic structure of the human prototype strain H of hepatitis C virus: C		
A;Reference number: A36814		
A;Accession: A36814		
A;Molecule type: genomic RNA		
A;Residues: 1-3011 <NC>		
A;Cross-references: UNIPROT:P27958; GB:M67463; NID:G329737; PIDN:AAA45534.1; PID:G329738		
R;Inchauspe, G.; Zebedee, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.		
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991		
A;Title: Genomic structure of the human prototype strain H of hepatitis C virus: comparati		
A;Reference number: A41546; MUID:92052256; PMID:1658800		
A;Contents: annotation		
A;Note: neither amino acid nor nucleotide sequence is given		
C;Superfamily: hepatitis C virus genome polyprotein		
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura		
F;1-115/Product: capsid protein C #status predicted <CPC>		
F;116-191/Product: envelope protein M #status predicted <EPW>		
F;192-389/Product: major envelope protein E #status predicted <MEE>		
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>		
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>		
F;1007-1615/Product: hepatitis C virus #status predicted <NS3>		
F;1230-1237/Region: nucleotide-binding motif A (P-loop)		
F;1312-1317/Region: nucleotide-binding motif B		
F;1316-1319/Region: DEXH motif		
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>		
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>		
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>		
F;196, 209, 234, 305, 325, 417, 423, 430, 446, 476, 532, 540, 556, 576, 623, 645, 1213, 1255, 2041, 2240, 23		
Query Match 99.1%; Score 1031; DB 1; Length 3011;		
Best Local Similarity 99.0%; Pred. No. 4.8e-77; Indels 0; Gaps 0;		
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;		
Qy	1	MSTNPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Db	1	MSTNPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Qy	61	RRQIPKARRPEGRTWAOQGYWPWLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
Db	61	RRQIPKARRPEGRTWAOQGYWPWLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
Qy	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNYYATGNLPGCSFSIFLLA 180
Db	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNYYATGNLPGCSFSIFLLA 180
Qy	181	LLSCLTVPASA 191
Db	181	LLSCLTVPASA 191

RESULT 3		
GNWVCJ		
genome polyprotein - hepatitis C virus (strain J)		
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural		
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5		
C;Species: hepatitis C virus		
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004		
A;Accession: A39253; PS0086		
R;Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimoto		
Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990		
A;Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients v		
A;Reference number: A39253; MUID:91088550; PMID:2175903		
A;Accession: A39253		
A;Molecule type: genomic RNA		
A;Residues: 1-3010 <KAT>		
A;Cross-references: UNIPROT:P26662; GB:D90208; NID:G221610; PIDN:BA414233.1; PID:G221611		
R;Kato, N.; Ohkoshi, S.; Shimotohno, K.		
Proc. Jpn. Acad. 65B, 219-223, 1989		
A;Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari		
A;Reference number: PS0085		
A;Accession: PS0086		
A;Molecule type: genomic RNA		
A;Residues: 2650-2707 <KA2>		
A;Experimental source: Japanese isolate		
C;Comment: The cleavage sites of this polyprotein have not been determined.		
C;Superfamily: hepatitis C virus genome polyprotein		
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine		
F;2-115/Product: capsid protein C #status predicted <CPC>		
F;116-191/Product: envelope protein M #status predicted <BPM>		
F;192-389/Product: major envelope protein E #status predicted <MEE>		
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>		
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>		
F;1007-1615/Product: hepatitis C virus #status predicted <NS3>		
F;1230-1237/Region: nucleotide-binding motif A (P-loop)		
F;1312-1317/Region: nucleotide-binding motif B		
F;1316-1319/Region: DEXH motif		
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>		
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>		
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>		
F;196, 209, 234, 305, 325, 417, 423, 430, 448, 532, 556, 576, 623, 645, 1213, 1255, 2041, 2077, 2240, 2;		
Query Match 98.1%; Score 1020; DB 1; Length 3010;		
Best Local Similarity 97.9%; Pred. No. 3.9e-76; Indels 0; Gaps 0;		
Matches 187; Conservative 2; Mismatches 2; Indels 0; Gaps 0;		
Qy	1	MSTNPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Db	1	MSTNPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Qy	61	RRQIPKARRPEGRTWAOQGYWPWLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
Db	61	RRQIPKARRPEGRTWAOQGYWPWLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
Qy	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNYYATGNLPGCSFSIFLLA 180
Db	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNYYATGNLPGCSFSIFLLA 180
Qy	181	LLSCLTVPASA 191
Db	181	LLSCLTVPASA 191
RESULT 4		
S40770		
genome polyprotein - hepatitis C virus		
N;Contains: capsid protein C; envelope protein M; hepatitis C virus (strain J)		
Protein NS4a; nonstructural protein NS4b; nonstructural protein NS5		
C;Species: hepatitis C virus		
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004		
A;Accession: S40770; PC1285		
R;Okamoto, H.		

submitted to the EMBL Data Library, March 1992

A;Reference number: S40770  
A;Accession: S40770  
A;Molecule type: genomic RNA  
A;Residues: 1-3011 <OK>  
A;Cross-references: UNIPROT:Q03463; EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g221586  
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990

A;Title: The 5'-terminal sequence of the hepatitis C virus genome.

A;Reference number: PC1284; MUID:91013116; PMID:2170712

A;Accession: PC1285

A;Molecule type: genomic RNA

A;Residues: 1-513 <OK>

A;Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512

A;Experimental source: isolate HC-J1

C;Superfamily: hepatitis C virus genome polypeptide

C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; p-loop; polypeptide; serine

F;2-115/Product: capsid protein C #status predicted <CPC>

F;116-191/Product: envelope protein M #status predicted <EPM>

F;192-389/Product: major envelope protein E #status predicted <MEE>

F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F;1007-1615/Product: hepatitis virus #status predicted <NS3>

F;1230-1237/Product: nucleotide-binding motif A (p-loop)

F;1312-1317/Product: nucleotide-binding motif B

F;1316-1319/Product: DEXH motif

F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>

F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4b>

F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match 97.9%; Score 1018; DB 1; Length 3011;  
Best Local Similarity 97.9%; Pred. No. 5.7e-76;  
Matches 187; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60

DB 1 MSTIPKQKTKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60

QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECGWAGWLLSPRGSPPSWGPTDPRRSNLTG 120

DB 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECGWAGWLLSPRGSPPSWGPTDPRRSNLTG 120

QY 121 KVDTLTGCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

DB 121 KVDTLTGCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191

DB 181 LLSCLTVPASA 191

RESULT 5

JQ1584

genome polypeptide - hepatitis C virus (strain U.K.) (fragment)

N;Contains: core protein C; envelope protein E1; envelope protein E2; nonstructural protein

C;Species: hepatitis C virus

C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004

C;Accession: JQ1584

R;Kumar, U.; Cheng, D.; Thomas, H.; Monjardino, J.

J. Gen. Virol. 73, 1521-1525, 1992

A;Title: Cloning and sequencing of the structural region and expression of putative core

A;Reference number: JQ1584; MUID:92300349; PMID:1318944

A;Accession: JQ1584

A;Molecule type: genomic RNA

A;Residues: 1-640 <KUM>

A;Cross-references: UNIPROT:Q68966; GB:X84079; NID:g643119; PIDN:CAA58888.1; PID:g643120

C;Superfamily: hepatitis C virus genome polypeptide

C;Keywords: core protein; envelope protein; glycoprotein; nonstructural protein; polypeptide

F;1-191/Product: core protein C #status predicted <CPC>

F;192-389/Product: envelope protein E1 #status predicted <EE1>

F;390-640/Product: envelope protein E2 and nonstructural protein NS1 #status predicted <NS1>

F;196, 209, 234, 305, 417, 430, 448, 476, 540, 556, 576, 623/Binding site: carbohydrate (Asn) (cova

Query Match 97.7%; Score 1016; DB 2; Length 640;  
Best Local Similarity 97.4%; Pred. No. 2e-76;  
Matches 186; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60

DB 1 MSTNPKPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60

QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECGWAGWLLSPRGSPPSWGPTDPRRSNLTG 120

DB 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECGWAGWLLSPRGSPPSWGPTDPRRSNLTG 120

QY 121 KVDTLTGCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

DB 121 KVDTLTGCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191

DB 181 LLSCLTVPASA 191

RESULT 6

S12707

genome polypeptide - hepatitis C virus (fragment)

N;Contains: core protein; envelope protein

C;Species: hepatitis C virus

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004

C;Accession: S12707

R;Takeuchi, K.; Kubo, Y.; Boonmar, S.; Watanabe, Y.; Katayama, T.; Choo, Q.L.; Kuo, G.;

Nucleic Acids Res. 18, 4626, 1990

A;Title: Nucleotide sequence of core and envelope genes of the hepatitis C virus genome

A;Reference number: S12707; MUID:90356432; PMID:2117749

A;Accession: S12707

A;Molecule type: genomic RNA

A;Residues: 1-441 <TAK>

A;Cross-references: UNIPROT:Q81776; EMBL:D00574; NID:g221656; PIDN:BAA00452.1; PID:g221656

C;Superfamily: hepatitis C virus genome polypeptide

C;Keywords: polypeptide

Query Match 97.6%; Score 1015; DB 2; Length 441;  
Best Local Similarity 97.4%; Pred. No. 1.7e-76;  
Matches 186; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60

DB 1 MSTNPKPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRGLGVRATKTSERSQPRG 60

QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECGWAGWLLSPRGSPPSWGPTDPRRSNLTG 120

DB 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECGWAGWLLSPRGSPPSWGPTDPRRSNLTG 120

QY 121 KVDTLTGCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

DB 121 KVDTLTGCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191

DB 181 LLSCLTVPASA 191

RESULT 7

S32740

polypeptide - hepatitis C virus (isolate Russian) (fragment)

N;Contains: capsid protein C; envelope protein M

C;Species: hepatitis C virus

C;Date: 19-Mar-1997 #sequence\_revision 05-Dec-1998 #text\_change 09-Jul-2004

C;Accession: S32740

R;Vassilev, V.B.; Vlazov, S.O.; Kotova, E.Y.; Nosikov, V.V.

submitted to the EMBL Data Library, April 1993

A;Description: Evidence of new HCV variant of European isolate in Russia.

A;Reference number: S32740

A;Accession: S32740

A;Molecule type: genomic RNA



A;Residues: 1-189 <VAS>  
A;Cross-references: UNIPROT:Q68873; EMBL:X71407  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: capsid protein; envelope protein; polyprotein  
F;1-115/Product: capsid protein C #status predicted <CPC>  
F;116-189/Product: envelope protein M #status predicted <EPM>  
  
Query Match 97.5%; Score 1014; DB 2; Length 189;  
Best Local Similarity 96.4%; Pred. No. 9.1e-77;  
Matches 186; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 MSTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
DB 1 MSTNPKQKNTNRRPDQVFPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
  
QY 61 RRQPIPKARRPEGRWTAQPCGYPWPLYGNECCGAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
DB 61 RRQPIPKARRPEGRWTAQPCGYPWPLYGNECCGAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
  
QY 121 KVDTLTTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
DB 121 KVDTLTTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
  
QY 181 LLSCLTTPA 189  
DB 181 LLSCLTTPA 189  
  
RESULT 8  
S18031  
genome polyprotein - hepatitis C virus (isolate JK2) (fragment)  
N;Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein  
C;Species: hepatitis C virus  
A;Variety: isolate JK2  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
R;Honda, M.; Kaneko, S.; Maeaishi, U.; Kobayashi, K.; Murakami, S.  
submitted to the EMBL Data Library, September 1991  
A;Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso  
A;Reference number: S18029  
A;Accession: S18031  
A;Molecule type: genomic RNA  
A;Residues: 1-782 <HON>  
A;Cross-references: UNIPROT:Q68950; EMBL:X61593  
A;Experimental source: isolate JK2  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural  
F;1-191/Product: core protein #status predicted <MAT1>  
F;192-383/Product: envelope protein 1 #status predicted <MAT2>  
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>  
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>  
  
Query Match 97.2%; Score 1011; DB 2; Length 782;  
Best Local Similarity 96.9%; Pred. No. 6.1e-76;  
Matches 185; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 MSTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
DB 1 MSTNPKQKNTNRRPDQVFPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
  
QY 61 RRQPIPKARRPEGRWTAQPCGYPWPLYGNECCGAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
DB 61 RRQPIPKARRPEGRWTAQPCGYPWPLYGNECCGAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
  
QY 121 KVDTLTTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
DB 121 KVDTLTTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
  
QY 181 LLSCLTTPA 191  
DB 181 LLSCLTTPA 191

RESULT 9  
PN0677  
hypothetical protein 787 - hepatitis C virus (fragment)  
C;Species: hepatitis C virus  
C;Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 09-Jul-2004  
C;Accession: PN0677  
R;Cho, S.H.; Yoon, J.I.; Chang, J.E.; Ahn, B.M.; Lee, C.H.; Lee, Y.I.  
Biochem. Biophys. Res. Commun. 196; 780-789, 1993  
A;Title: Genomic typing of hepatitis C viruses from Korean patients: Implications of gen  
A;Reference number: PN0677; MUID:94059104; PMID:8240354  
A;Accession: PN0677  
A;Molecule type: mRNA  
A;Residues: 1-787 <CHO>  
A;Cross-references: UNIPROT:Q08244; GB:L20498; NID:gl381031; PIDN:AAB02608.1; PID:gl38103  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: glycoprotein; nonstructural protein  
F;196,209,234,250,305,325,421,427,452,536,544,560,580,627,649/Binding site: carbohydrate  
  
Query Match 97.2%; Score 1011; DB 2; Length 787;  
Best Local Similarity 96.9%; Pred. No. 6.1e-76;  
Matches 185; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 MSTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
DB 1 MSTNPKQKNTNRRPDQVFPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
  
QY 61 RRQPIPKARRPEGRWTAQPCGYPWPLYGNECCGAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
DB 61 RRQPIPKARRPEGRWTAQPCGYPWPLYGNECCGAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
  
QY 121 KVDTLTTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
DB 121 KVDTLTTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
  
QY 181 LLSCLTTPA 191  
DB 181 LLSCLTTPA 191  
  
RESULT 10  
GNWVTC  
genome polyprotein - hepatitis C virus  
N;Contains: capsid protein C; envelope protein M; hepatitis C virus genome polyprotein  
C;Species: hepatitis C virus  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C;Accession: A38465  
R;Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.; J  
J. Virol. 65; 1105-1113, 1991  
A;Title: Structure and organization of the hepatitis C virus genome isolated from human  
A;Reference number: A38465; MUID:91140698; PMID:1847440  
A;Accession: A38465  
A;Molecule type: genomic RNA  
A;Residues: 1-3010 <TAK>  
A;Cross-references: EMBL:M58335; NID:g329770; PIDN:AAA72945.1; PID:g3297;  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural  
F;2-115/Product: capsid protein C #status predicted <CPC>  
F;116-191/Product: envelope protein M #status predicted <EPM>  
F;192-389/Product: major envelope protein NS1 #status predicted <MEE>  
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F;1007-1615/Product: hepatitis C virus genome polyprotein NS3 #status predicted <NS3>  
F;1230-1237/Region: nucleotide-binding motif A (P-loop)  
F;1312-1317/Region: nucleotide-binding motif B  
F;1316-1319/Region: DEH motif  
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>  
F;1863-2013/Product: nonstructural protein NS5 #status predicted <N4B>  
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>  
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,224  
  
Query Match 97.2%; Score 1011; DB 1; Length 3010;  
Best Local Similarity 97.4%; Pred. No. 2.2e-75;





Proc. Natl. Acad. Sci. U.S.A. 89, 3190-3194, 1992  
A>Title: Interaction of immune sera with synthetic peptides corresponding to the structure of hepatitis C virus genome polyprotein  
A;Reference number: A44150; MUID:92228749; PMID:1373489  
A;Accession: A44150  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: genomic RNA  
A;Residues: 1-513 <CHI>  
A;Cross-references: UNIPROT:Q91FE5; UNIPROT:Q34579; UNIPROT:Q36610; UNIPROT:Q9ELS8; UNIPROT:Q9ELS9  
C;Superfamily: hepatitis C virus genome polyprotein

Query Match 96.0%; Score 998; DB 2; Length 513;  
Best Local Similarity 97.9%; Pred. No. 4.9e-75;  
Matches 188; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 MSTNPKQKKNKNTNRRPQDVKPPGGQIVGGYLLPRRGLGVRAATKTSERSQPRG 60  
Db 1 MSTNPKQIKTKNTNRRPQDVKPPGGQIVGGYLLPRRGLGVRAATKTSERSQPRG 60

Qy 61 RROPIKARPEGRRTWAQPGYPLVGNCGWAGWLLSPRGR-PSWGPTDPRRSRNL 119  
Db 61 RROPIKARPEGRRTWAQPGYPLVGNCGWAGWLLSPRGRPSWG-TDPRRSRNL 119

Qy 120 GKVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLL 179  
Db 120 GKVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLL 179

Qy 180 ALLSCLTVPASA 191  
Db 180 ALLSCLTVPASA 191

RESULT 15  
A45573  
genome polyprotein - hepatitis C virus (strain JT)  
N;Contains: capsid protein C; envelope protein M; hepatitis C virus (strain JT) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: hepatitis C virus  
C;Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C;Accession: A45573  
R;Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata, Y. Virus Res. 23, 39-53, 1992  
A>Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier: a reference number: A45573; MUID:92295714; PMID:1318627  
A;Accession: A45573  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-3010 <TAN>  
A;Cross-references: UNIPROT:Q00269; GB:D11168; GB:D01171; NID:G221612; PIDN:BAA01943.1; A;Experimental source: HCV-JT  
A;Note: sequence extracted from NCBI backbone (NCBI:106206, NCBI:106207)  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serine  
F;2-115/Product: capsid protein C #status predicted <CPC>  
F;116-191/Product: envelope protein M #status predicted <EPM>  
F;192-389/Product: major envelope protein E #status predicted <MEE>  
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F;1007-1615/Product: hepatitis C virus genome polyprotein  
F;1230-1237/Region: nucleotide-binding motif A (P-loop)  
F;1312-1317/Region: nucleotide-binding motif B  
F;1316-1319/Region: DEXH motif  
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>  
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>  
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match 96.0%; Score 998; DB 1; Length 3010;  
Best Local Similarity 95.8%; Pred. No. 2.6e-74;  
Matches 183; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPQDVKPPGGQIVGGYLLPRRGLGVRAATKTSERSQPRG 60  
Db 1 MSTNPKQIKTKNTNRRPQDVKPPGGQIVGGYLLPRRGLGVRAATKTSERSQPRG 60

Qy 61 RROPIKARPEGRRTWAQPGYPLVGNCGWAGWLLSPRGRPSWGPTDPRRSRNLG 120  
Db 61 RROPIKARPEGRRTWAQPGYPLVGNCGWAGWLLSPRGRPSWGPTDPRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

Search completed: August 19, 2005, 12:44:07  
Job time : 25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 12:43:21 ; Search time 107 Seconds  
(without alignments)

698.999 Million cell updates/sec

Title: US-10-770-117-4

Perfect score: 1040

Sequence: 1 MSTNPKPQKKKNTNRPPQ.....CSPSIFLLALLSCLTVPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1040	100.0	191	18	US-10-770-117-4
2	1040	100.0	3011	9	US-09-916-359-2
3	1040	100.0	3011	16	US-10-445-724-2
4	1036	99.6	190	14	US-10-268-562-1
5	1035	99.5	190	15	US-10-450-649-7
6	1035	99.5	191	18	US-10-770-117-2
7	1031	99.1	249	15	US-10-365-620-54
8	1031	99.1	249	17	US-10-912-969-56
9	1031	99.1	450	15	US-10-651-165-179
10	1031	99.1	450	15	US-10-651-165-180
11	1031	99.1	473	15	US-10-365-620-56

12	1031	99.1	473	17	US-10-912-969-58	Sequence 58, Appl
13	1031	99.1	473	17	US-10-913-171-39	Sequence 39, Appl
14	1031	99.1	3011	9	US-09-742-659-4	Sequence 4, Appl
15	1031	99.1	3011	9	US-09-952-572-9	Sequence 9, Appl
16	1031	99.1	3011	9	US-09-929-955-1	Sequence 1, Appl
17	1031	99.1	3011	9	US-09-747-419-20	Sequence 20, Appl
18	1031	99.1	3011	10	US-09-891-894-3	Sequence 3, Appl
19	1031	99.1	3011	13	US-10-104-966-1	Sequence 1, Appl
20	1031	99.1	3011	14	US-10-259-275-20	Sequence 20, Appl
21	1031	99.1	3011	14	US-10-184-150-3	Sequence 3, Appl
22	1031	99.1	3011	15	US-10-328-997-3	Sequence 3, Appl
23	1031	99.1	3011	15	US-10-189-359-14	Sequence 14, Appl
24	1031	99.1	3011	15	US-10-296-734-406	Sequence 406, Appl
25	1031	99.1	3011	15	US-10-719-619-1	Sequence 1, Appl
26	1031	99.1	3011	16	US-10-817-591-1	Sequence 1, Appl
27	1031	99.1	3011	20	US-11-006-313-20	Sequence 20, Appl
28	1031	99.1	3012	9	US-09-238-076-2	Sequence 2, Appl
29	1031	99.1	3012	10	US-09-995-937-2	Sequence 2, Appl
30	1031	99.1	3012	10	US-09-917-563-2	Sequence 2, Appl
31	1027	98.8	3011	9	US-09-238-076-20	Sequence 20, Appl
32	1027	98.8	3011	10	US-09-995-937-20	Sequence 20, Appl
33	1027	98.8	3011	10	US-09-917-563-20	Sequence 20, Appl
34	1020	98.1	450	15	US-10-651-165-187	Sequence 187, Appl
35	1018	97.9	450	15	US-10-651-165-181	Sequence 181, Appl
36	1018	97.9	2894	9	US-09-941-611-23	Sequence 23, Appl
37	1018	97.9	2894	14	US-10-044-995-23	Sequence 23, Appl
38	1018	97.9	2894	17	US-10-822-871-23	Sequence 23, Appl
39	1016	97.7	450	15	US-10-651-165-190	Sequence 190, Appl
40	1016	97.7	3011	14	US-10-232-643-6	Sequence 6, Appl
41	1012	97.3	809	9	US-09-973-025-50	Sequence 50, Appl
42	1012	97.3	809	10	US-09-899-303-50	Sequence 50, Appl
43	1012	97.3	809	10	US-09-995-808-50	Sequence 50, Appl
44	1012	97.3	809	10	US-09-995-860-50	Sequence 50, Appl
45	1012	97.3	809	10	US-09-995-791-50	Sequence 50, Appl

ALIGNMENTS

RESULT 1

- US-10-770-117-4
- Sequence 4, Application US/10770117
- Publication No. US20050129705A1
- GENERAL INFORMATION:
- APPLICANT: The Government of the United States, as represented by the
- APPLICANT: Secretary, Dept. of
- APPLICANT: Health and Human Services
- APPLICANT: Berzofsky, Jay A.
- APPLICANT: Sarobe, Pablo
- APPLICANT: Pendleton, C. David
- APPLICANT: Feinstein, Stephen M.
- APPLICANT: Arichi, Tatsumi
- APPLICANT: Major, Marian E.
- TITLE OF INVENTION: MODIFIED HCV PEPTIDE VACCINES
- FILE REFERENCE: 14014.0347/P
- CURRENT APPLICATION NUMBER: US/10770,117
- CURRENT FILING DATE: 2004-02-02
- PRIOR APPLICATION NUMBER: US/09/763,260
- PRIOR FILING DATE: 2001-10-19
- PRIOR APPLICATION NUMBER: 60/097,446
- PRIOR FILING DATE: 1998-08-21
- NUMBER OF SEQ ID NOS: 8
- SOFTWARE: FastSeq for Windows Version 3.0
- SEQ ID NO 4
- LENGTH: 191
- TYPE: PRT
- ORGANISM: Artificial Sequence
- FEATURE:
- OTHER INFORMATION: Description of Artificial Sequence:/Note =
- OTHER INFORMATION: Synthetic construct
- US-10-770-117-4

Query Match 100.0%; Score 1040; DB 18; Length 191;

Best Local Similarity 100.0%; Pred. No. 5.7e-81;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MSTNPKQKKKNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
Db 1 MSTNPKQKKKNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
  
Qy 61 RRQIPKARRPEGRWTAQPGYWPPLYNCGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Db 61 RRQIPKARRPEGRWTAQPGYWPPLYNCGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
  
Qy 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNATGNLPGCSFSIFLLA 180  
Db 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNATGNLPGCSFSIFLLA 180  
  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

RESULT 2

US-09-916-359-2  
; Sequence 2, Application US/09916359  
; Patent No. US20020034734A1  
; GENERAL INFORMATION:  
; APPLICANT: Veronique Barban  
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR  
; TITLE OF INVENTION: TREATING C HEPATITIS  
; FILE REFERENCE: PMCF97-03A  
; CURRENT APPLICATION NUMBER: US/09/916,359  
; CURRENT FILING DATE: 2001-07-26  
; PRIOR FILING DATE: 09/388,874  
; PRIOR FILING DATE: 1999-09-02  
; PRIOR FILING DATE: 1997-03-06  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 3011  
; TYPE: PRT  
; ORGANISM: Virus  
US-09-916-359-2

Query Match 100.0%; Score 1040; DB 9; Length 3011;  
Best Local Similarity 100.0%; Pred. No. 1.1e-79;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MSTNPKQKKKNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
Db 1 MSTNPKQKKKNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
  
Qy 61 RRQIPKARRPEGRWTAQPGYWPPLYNCGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Db 61 RRQIPKARRPEGRWTAQPGYWPPLYNCGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
  
Qy 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNATGNLPGCSFSIFLLA 180  
Db 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNATGNLPGCSFSIFLLA 180  
  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

RESULT 3

US-10-445-724-2  
; Sequence 2, Application US/10445724  
; Publication No. US20040101829A1  
; GENERAL INFORMATION:  
; APPLICANT: STAPLETON, JACK T.  
; APPLICANT: WUENSCHMANN, SABINA  
; TITLE OF INVENTION: A PROTEIN THAT INTERACTS WITH LIPIDS AND  
; TITLE OF INVENTION: METHODS FOR TREATING HYPERLIPIDEMIA

; FILE REFERENCE: IOWA:045US  
; CURRENT APPLICATION NUMBER: US/10/445,724  
; CURRENT FILING DATE: 2003-05-27  
; PRIOR APPLICATION NUMBER: 60/392,158  
; PRIOR FILING DATE: 2003-05-27  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 3011  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-10-445-724-2

Query Match 100.0%; Score 1040; DB 16; Length 3011;  
Best Local Similarity 100.0%; Pred. No. 1.1e-79;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MSTNPKQKKKNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
Db 1 MSTNPKQKKKNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
  
Qy 61 RRQIPKARRPEGRWTAQPGYWPPLYNCGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Db 61 RRQIPKARRPEGRWTAQPGYWPPLYNCGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
  
Qy 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNATGNLPGCSFSIFLLA 180  
Db 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNATGNLPGCSFSIFLLA 180  
  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

RESULT 4

US-10-268-562-1  
; Sequence 1, Application US/10268562  
; Publication No. US20030108563A1  
; GENERAL INFORMATION:  
; APPLICANT: Ocho-Clinical Diagnostics, Inc.  
; TITLE OF INVENTION: Reagents for the simultaneous detection of HCV core antigens and  
; TITLE OF INVENTION: antibodies  
; FILE REFERENCE: CDS0287  
; CURRENT APPLICATION NUMBER: US/10/268,562  
; CURRENT FILING DATE: 2002-10-10  
; PRIOR APPLICATION NUMBER: US 60/347,943  
; PRIOR FILING DATE: 2001-11-07  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 190  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-10-268-562-1

Query Match 99.6%; Score 1036; DB 14; Length 190;  
Best Local Similarity 100.0%; Pred. No. 1.2e-80;  
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 MSTNPKQKKKNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
Db 1 MSTNPKQKKKNTNRRPQDVKFGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
  
Qy 61 RRQIPKARRPEGRWTAQPGYWPPLYNCGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Db 61 RRQIPKARRPEGRWTAQPGYWPPLYNCGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
  
Qy 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNATGNLPGCSFSIFLLA 180  
Db 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRLVDGVNATGNLPGCSFSIFLLA 180  
  
Qy 181 LLSCLTVPASA 190  
Db 181 LLSCLTVPASA 190

```
Db      181  LLSCLTVPAS 190

RESULT 5
US-10-450-649-7
; Sequence 7, Application US/10450649
; Publication No. US20040052818A1
; GENERAL INFORMATION:
; APPLICANT: Heinz, Franz X.
; APPLICANT: Mandl, Christian
; TITLE OF INVENTION: ATTENUATED LIVE VACCINE
; FILE REFERENCE: U 014666-0
; CURRENT APPLICATION NUMBER: US/10/450,649
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: PCT/AT02/00046
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: A 272/2001 AT
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Hepatitis C Virus 1
US-10-450-649-7

Query Match      99.5%; Score 1035; DB 15; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.5e-80;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2  STNPKPKKKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVTRATKTSERSQPRGR 61
Db      1  STNPKPKKKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVTRATKTSERSQPRGR 60
Qy      62  RQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGRSPSWGPTDPRRRSRNLGK 121
Db      61  RQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGRSPSWGPTDPRRRSRNLGK 120
Qy      122  VIDTLTCGFADLMGYIPLVGAPLGGARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 181
Db      121  VIDTLTCGFADLMGYIPLVGAPLGGARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Qy      182  LLSCLTVPASA 191
Db      181  LLSCLTVPASA 190

RESULT 6
US-10-770-117-2
; Sequence 2, Application US/10770117
; Publication No. US20050129705A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States, as represented by the
; APPLICANT: Secretary, Dept. of
; APPLICANT: Health and Human Services
; APPLICANT: Berzofsky, Jay A.
; APPLICANT: Sarobe, Pablo
; APPLICANT: Pendleton, C. David
; APPLICANT: Feinstein, Stephen M.
; APPLICANT: Arichi, Tatsumi
; APPLICANT: Major, Marian E.
; TITLE OF INVENTION: MODIFIED HCV PEPTIDE VACCINES
; FILE REFERENCE: 14014.0347/P
; CURRENT APPLICATION NUMBER: US/10/770,117
; CURRENT FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: US/09/763,260
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 60/097,446
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 191
; TYPE: PRT
; ORGANISM: ORF of HCV Core Protein
US-10-365-620-54
; Sequence 54, Application US/10365620
; Publication No. US20040001853A1
; GENERAL INFORMATION:
; APPLICANT: George, Rajan
; APPLICANT: Tyrrell, Lorne
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 656.0016
; CURRENT APPLICATION NUMBER: US/10/365,620
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US60/423,578
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/390,564
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54
; LENGTH: 249
; TYPE: PRT
; ORGANISM: ORF of HCV Core Protein
US-10-365-620-54

Query Match      99.1%; Score 1031; DB 15; Length 249;
Best Local Similarity 99.0%; Pred. No. 4.4e-80;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy      1  MSTNPKPKKKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVTRATKTSERSQPRG 60
Db      31  MSTNPKPKKKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVTRATKTSERSQPRG 90
Qy      61  RQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
Db      91  RQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 150
Qy      121  KVDTTLTCGFADLMGYIPLVGAPLGGARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db      151  KVDTTLTCGFADLMGYIPLVGAPLGGARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 210
Qy      181  LLSCLTVPASA 191
Db      211  LLSCLTVPASA 221

RESULT 7
US-10-365-620-54
; Sequence 54, Application US/10365620
; Publication No. US20040001853A1
; GENERAL INFORMATION:
; APPLICANT: George, Rajan
; APPLICANT: Tyrrell, Lorne
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 656.0016
; CURRENT APPLICATION NUMBER: US/10/365,620
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US60/423,578
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/390,564
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54
; LENGTH: 249
; TYPE: PRT
; ORGANISM: ORF of HCV Core Protein
US-10-365-620-54

Query Match      99.5%; Score 1035; DB 18; Length 191;
Best Local Similarity 99.5%; Pred. No. 1.5e-80;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1  MSTNPKPKKKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVTRATKTSERSQPRG 60
Db      1  MSTNPKPKKKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVTRATKTSERSQPRG 60
Qy      61  RQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
Db      61  RQPIPKARRPEGRRTWAQPGYPMPLYGNECGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
Qy      121  KVDTTLTCGFADLMGYIPLVGAPLGGARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db      121  KVDTTLTCGFADLMGYIPLVGAPLGGARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Qy      181  LLSCLTVPASA 191
Db      181  LLSCLTVPASA 191
```

```
RESULT 8
US-10-912-969-56
; Sequence 56, Application US/10912969
; Publication No. US20050013828A1
; GENERAL INFORMATION:
; APPLICANT: Virex Research, Inc.
; APPLICANT: George, Rajan
; APPLICANT: Tyrrell, Lorne
; APPLICANT: Noujaim, Antoine
; APPLICANT: Wang, Dakun
; APPLICANT: Ma, Allan
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 17506-007001
; CURRENT APPLICATION NUMBER: US/10/912,969
; CURRENT FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: US 60/390,564
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/423,578
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 10/365,620
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: PCT/IB04/00373
; PRIOR FILING DATE: 2004-02-14
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-912-969-56

Query Match          99.1%; Score 1031; DB 17; Length 249;
Best Local Similarity 99.0%; Pred. No. 4.4e-80;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKNKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
Db 31 MSTNPKPQRTKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 90
QY 61 RRQPIKARRPEGRWTWAOQGYPMPLYGNECGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 120
Db 91 RRQPIKARRPEGRWTWAOQGYPMPLYGNECGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 150
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180
Db 151 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 210
QY 181 LLSCLTVPASA 191
Db 211 LLSCLTVPASA 221

Query Match          99.1%; Score 1031; DB 17; Length 249;
Best Local Similarity 99.0%; Pred. No. 4.4e-80;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKNKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
Db 31 MSTNPKPQRTKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 90
QY 61 RRQPIKARRPEGRWTWAOQGYPMPLYGNECGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 120
Db 91 RRQPIKARRPEGRWTWAOQGYPMPLYGNECGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 150
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180
Db 151 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 210
QY 181 LLSCLTVPASA 191
Db 211 LLSCLTVPASA 221

RESULT 9
US-10-651-165-179
; Sequence 179, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 180
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-179

Query Match          99.1%; Score 1031; DB 15; Length 450;
Best Local Similarity 99.0%; Pred. No. 8.3e-80;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKNKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
Db 1 MSTNPKPQRTKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
QY 61 RRQPIKARRPEGRWTWAOQGYPMPLYGNECGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 120
Db 61 RRQPIKARRPEGRWTWAOQGYPMPLYGNECGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

Query Match          99.1%; Score 1031; DB 15; Length 450;
Best Local Similarity 99.0%; Pred. No. 8.3e-80;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKNKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
Db 1 MSTNPKPQRTKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
QY 61 RRQPIKARRPEGRWTWAOQGYPMPLYGNECGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 120
Db 61 RRQPIKARRPEGRWTWAOQGYPMPLYGNECGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVEDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191
```

```
RESULT 11
US-10-365-620-56
; Sequence 56, Application US/10365620
; Publication No. US20040001853A1
; GENERAL INFORMATION:
; APPLICANT: George, Rajan
; APPLICANT: Tyrell, Lorne
; APPLICANT: No. US20040001853A1jaim, Antoine
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 656.0016
; CURRENT APPLICATION NUMBER: US/10/365,620
; CURRENT FILING DATE: 2003-02-13
; PRIOR FILING DATE: 2003-11-05
; PRIOR FILING DATE: 2002-06-20
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 473
; TYPE: PRT
; ORGANISM: ORF of HCV Core-TBD protein
US-10-365-620-56

Query Match          99.1%; Score 1031; DB 15; Length 473;
Best Local Similarity 99.0%; Pred. No. 8.8e-80;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
DB 31 MSTNPKPQRTKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 90
QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNEGCWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
DB 91 RRQPIPKARRPEGRRTWAQPGYPWPLYGNEGCWAGWLLSPRGRSPSWGPTDPRRRSRNLG 150
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
DB 151 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 210
QY 181 LLSCLTVPASA 191
DB 211 LLSCLTVPASA 221

RESULT 12
US-10-912-969-58
; Sequence 58, Application US/10912969
; Publication No. US20050013828A1
; GENERAL INFORMATION:
; APPLICANT: Virexx Research, Inc.
; APPLICANT: George, Rajan
; APPLICANT: Tyrell, Lorne
; APPLICANT: Noujaim, Antoine
; APPLICANT: Wang, Dakun
; APPLICANT: Ma, Allan
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 17506-007001
; CURRENT APPLICATION NUMBER: US/10/912,969
; CURRENT FILING DATE: 2004-08-05
; PRIOR FILING DATE: 2002-06-20
; PRIOR FILING DATE: 2002-11-05
; PRIOR FILING DATE: 2003-02-13
; PRIOR FILING DATE: 2004-02-14
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58
```

```
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-912-969-58

Query Match          99.1%; Score 1031; DB 17; Length 473;
Best Local Similarity 99.0%; Pred. No. 8.8e-80;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
DB 31 MSTNPKPQRTKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 90
QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNEGCWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
DB 91 RRQPIPKARRPEGRRTWAQPGYPWPLYGNEGCWAGWLLSPRGRSPSWGPTDPRRRSRNLG 150
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
DB 151 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 210
QY 181 LLSCLTVPASA 191
DB 211 LLSCLTVPASA 221

RESULT 13
US-10-913-171-39
; Sequence 39, Application US/10913171
; Publication No. US20050031628A1
; GENERAL INFORMATION:
; APPLICANT: Virexx Research, Inc.
; APPLICANT: George, Rajan
; APPLICANT: Tyrell, Lorne
; APPLICANT: Noujaim, Antoine
; APPLICANT: Wang, Dakun
; APPLICANT: Ma, Allan
; TITLE OF INVENTION: CHIMERIC ANTIGENS FOR BREAKING HOST TOLERANCE TO FOREIGN ANTIGENS
; FILE REFERENCE: 17506-006001
; CURRENT APPLICATION NUMBER: US/10/913,171
; CURRENT FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: US 60/493,449
; PRIOR FILING DATE: 2004-08-08
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-913-171-39

Query Match          99.1%; Score 1031; DB 17; Length 473;
Best Local Similarity 99.0%; Pred. No. 8.8e-80;
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
DB 31 MSTNPKPQRTKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 90
QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNEGCWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
DB 91 RRQPIPKARRPEGRRTWAQPGYPWPLYGNEGCWAGWLLSPRGRSPSWGPTDPRRRSRNLG 150
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
DB 151 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 210
QY 181 LLSCLTVPASA 191
DB 211 LLSCLTVPASA 221
```

Db	211	LLSCLTVPASA 221	
RESULT 14			
US-09-742-659-4			
; Sequence 4, Application US/09742659			
; Patent No. US20010034019A1			
; GENERAL INFORMATION:			
; APPLICANT: Hong, Zhi			
; APPLICANT: Butkiewicz, Nancy J.			
; APPLICANT: Zhong, Weidong			
; APPLICANT: Ingravallo, Paul			
; APPLICANT: Wright-Minogue, Jacquelyn			
; APPLICANT: Lau, Johnson Y.			
; APPLICANT: Lemon, Stanley M.			
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses			
; FILE REFERENCE: ID01116			
; CURRENT APPLICATION NUMBER: US/09/742,659			
; CURRENT FILING DATE: 2000-12-21			
; PRIOR APPLICATION NUMBER: US 60/171,469			
; PRIOR FILING DATE: 1999-12-22			
; NUMBER OF SEQ ID NOS: 16			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 4			
; LENGTH: 3011			
; TYPE: PRT			
; ORGANISM: Hepatitis C virus			
US-09-742-659-4			
Query Match			
Best Local Similarity 99.1%; Score 1031; DB 9; Length 3011;			
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	MSTNPKPQKKNKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60	
Db	1	MSTNPKPQKTKKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60	
Qy	61	RRQPIPKARPEGRRTWAQGYWPPLYGNECCGAGWLLSPRGSRPSWGPTDPRRRSRNLG 120	
Db	61	RRQPIPKARPEGRRTWAQGYWPPLYGNECCGAGWLLSPRGSRPSWGPTDPRRRSRNLG 120	
Qy	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180	
Db	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180	
Qy	181	LLSCLTVPASA 191	
Db	181	LLSCLTVPASA 191	
Search completed: August 19, 2005, 13:01:51			
Job time : 108 secs			

---

Db	211	LLSCLTVPASA 221	
RESULT 15			
US-09-952-572-9			
; Sequence 9, Application US/09952572			
; Patent No. US20020119495A1			
; GENERAL INFORMATION:			
; APPLICANT: HAWAII BIOTECHNOLOGY GROUP, Inc.			
; APPLICANT: NAKANO, Eileen			
; APPLICANT: HUMPHREYS, Tom			
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITION OF HEPATITIS C			
; FILE REFERENCE: HAWBIO1100			
; CURRENT APPLICATION NUMBER: US/09/952,572			
; CURRENT FILING DATE: 2001-09-13			
; PRIOR APPLICATION NUMBER: US 60/230,927			
; PRIOR FILING DATE: 2000-09-13			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 9			
; LENGTH: 3011			
; TYPE: PRT			
; ORGANISM: Hepatitis C Virus			
US-09-952-572-9			
Query Match			
Best Local Similarity 99.0%; Pred. No. 6.2e-79;			
Matches 189; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	MSTNPKPQKKNKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60	
Db	1	MSTNPKPQKTKKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60	
Qy	61	RRQPIPKARPEGRRTWAQGYWPPLYGNECCGAGWLLSPRGSRPSWGPTDPRRRSRNLG 120	
Db	61	RRQPIPKARPEGRRTWAQGYWPPLYGNECCGAGWLLSPRGSRPSWGPTDPRRRSRNLG 120	
Qy	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180	
Db	121	KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180	
Qy	181	LLSCLTVPASA 191	
Db	181	LLSCLTVPASA 191	



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 12:34:55 ; Search time 20 Seconds  
(without alignments)  
712.899 Million cell updates/sec

Title: US-10-770-117-4

Perfect score: 1040

Sequence: 1 MSTNPKQKKNKNTNRPO.....CSPSIFLLALLSCLTVPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgm2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgm2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgm2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgm2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgm2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgm2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040	100.0	191	4	US-09-763-260-4
2	1040	100.0	3011	1	US-08-440-103-36
3	1040	100.0	3011	1	US-08-440-542-36
4	1040	100.0	3011	1	US-07-910-760-10
5	1040	100.0	3011	1	US-08-440-519-10
6	1040	100.0	3011	1	US-08-231-368-36
7	1040	100.0	3011	1	US-08-440-210-36
8	1040	100.0	3011	3	US-09-388-874-2
9	1040	100.0	3011	3	US-09-046-604-36
10	1040	100.0	3011	3	US-08-440-549-10
11	1040	100.0	3011	3	US-08-850-328-1
12	1040	100.0	3011	4	US-09-916-359-2
13	1036	99.6	215	1	US-07-681-703B-12
14	1036	99.6	215	1	US-07-681-703B-14
15	1036	99.6	215	5	PCT-US91-02370-12
16	1036	99.6	215	5	PCT-US91-02370-14
17	1036	99.6	217	2	US-08-407-410B-12
18	1036	99.6	217	2	US-08-407-410B-14
19	1036	99.6	217	2	US-08-485-500-12
20	1036	99.6	217	2	US-08-485-500-14
21	1035	99.5	191	4	US-09-763-260-2
22	1034	99.4	2995	3	US-08-444-818-138
23	1031	99.1	191	2	US-08-290-665A-156
24	1031	99.1	191	2	US-08-290-665A-157
25	1031	99.1	191	2	US-08-290-665A-158
26	1031	99.1	191	2	US-08-290-665A-159
27	1031	99.1	191	2	US-08-290-665A-160

28 1031 99.1 191 3 US-08-380-160-3 Sequence 3, Appli  
29 1031 99.1 191 5 PCT-US95-10398-156 Sequence 156, App  
30 1031 99.1 191 5 PCT-US95-10398-157 Sequence 157, App  
31 1031 99.1 191 5 PCT-US95-10398-158 Sequence 158, App  
32 1031 99.1 191 5 PCT-US95-10398-159 Sequence 159, App  
33 1031 99.1 191 5 PCT-US95-10398-160 Sequence 160, App  
34 1031 99.1 191 5 PCT-US95-10398-160 Sequence 179, App  
35 1031 99.1 450 4 US-08-635-886C-179 Sequence 180, App  
36 1031 99.1 450 4 US-08-635-886C-180 Sequence 179, App  
37 1031 99.1 450 4 US-08-974-690C-179 Sequence 180, App  
38 1031 99.1 967 1 US-08-188-281B-13 Sequence 13, Appl  
39 1031 99.1 967 5 PCT-US94-07280-13 Sequence 13, Appl  
40 1031 99.1 967 5 PCT-US95-01087-13 Sequence 13, Appl  
41 1031 99.1 1648 1 US-08-188-281B-12 Sequence 12, Appl  
42 1031 99.1 1648 5 PCT-US94-07280-12 Sequence 12, Appl  
43 1031 99.1 1648 5 PCT-US95-01087-12 Sequence 12, Appl  
44 1031 99.1 3011 1 US-08-188-281B-1 Sequence 1, Appl  
45 1031 99.1 3011 1 US-08-453-552-1 Sequence 1, Appl

#### ALIGNMENTS

RESULT 1  
US-09-763-260-4  
; Sequence 4, Application US/09763260  
; Patent No. 6685944  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States, as represented by the  
; APPLICANT: Secretary, Dept. of  
; APPLICANT: Health and Human Services  
; APPLICANT: Berzofsky, Jay A.  
; APPLICANT: Sarobe, Pablo  
; APPLICANT: Pendleton, C. David  
; APPLICANT: Feinstein, Stephen M.  
; APPLICANT: Arichi, Tatsumi  
; APPLICANT: Major, Marian E.  
; TITLE OF INVENTION: MODIFIED HCV PEPTIDE VACCINES  
; FILE REFERENCE: 14014.0347/P  
; CURRENT APPLICATION NUMBER: US/09/763,260  
; CURRENT FILING DATE: 2001-10-19  
; PRIOR APPLICATION NUMBER: 60/097,446  
; PRIOR FILING DATE: 1998-08-21  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/No. 6685944e =  
; OTHER INFORMATION: synthetic construct

US-09-763-260-4  
Query Match 100.0%; Score 1040; DB 4; Length 191;  
Best Local Similarity 100.0%; Pred. No. 4.5e-94;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MSTNPKQKKNKNTNRPODVKFGGQIVGGVYLLPRGPRGLGVATRTKTSERSOPRG 60  
Db 1 MSTNPKQKKNKNTNRPODVKFGGQIVGGVYLLPRGPRGLGVATRTKTSERSOPRG 60  
Qy 61 RRQIPKARPEGRWAOQGYPMPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
Db 61 RRQIPKARPEGRWAOQGYPMPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
Qy 121 KVIDTLTCGFADLMGYIPLVGAFLGGAARALAHGVRVLEQGVNATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPLVGAFLGGAARALAHGVRVLEQGVNATGNLPGCSFSIFLLA 180  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

```
RESULT 2
US-08-440-103-36
; Sequence 36, Application US/08440103
; Patent No. 5670152
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,103
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-440-103-36

Query Match 100.0%; Score 1040; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKQKKKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db 1 MSTNPKQKKKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
QY 61 RRQPIKARRPEGRTPAQGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
Db 61 RRQPIKARRPEGRTPAQGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSLCLTVPASA 191
Db 181 LLSLCLTVPASA 191

RESULT 3
US-08-440-542-36
; Sequence 36, Application US/08440542
; Patent No. 5670153
```

```
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,542
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-440-542-36

Query Match 100.0%; Score 1040; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKQKKKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db 1 MSTNPKQKKKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
QY 61 RRQPIKARRPEGRTPAQGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
Db 61 RRQPIKARRPEGRTPAQGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSLCLTVPASA 191
Db 181 LLSLCLTVPASA 191

RESULT 4
US-07-910-760-10
; Sequence 10, Application US/07910760
; Patent No. 5683864
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Choo, Qui-Liam
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
```

```
;
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/910,760
; FILING DATE: 07-JUL-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-910-760-10

Query Match 100.0%; Score 1040; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRWTAQPGYPWLYGNEGCGWAGWLLSPRGRSPSGWPTDPRRRSRNLG 120
Db 61 RRQIPKARRPEGRWTAQPGYPWLYGNEGCGWAGWLLSPRGRSPSGWPTDPRRRSRNLG 120

Qy 121 KVDTLTGCFADLMGYIPLVGAPLGGAARALAHGVRVLEGVNYATGNLPGCSFSIFLLA 180
Db 121 KVDTLTGCFADLMGYIPLVGAPLGGAARALAHGVRVLEGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 5
US-08-440-519-10
; Sequence 10, Application US/08440519
; Patent No. 5712087
; GENERAL INFORMATION:
; APPLICANT: Houghton, Michael
; APPLICANT: Kuo, George
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: P.O. Box 8097 (Int. Prop. R-440)
; CITY: Emeryville
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; Antibodies
```

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,519
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-440-519-10

Query Match 100.0%; Score 1040; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRWTAQPGYPWLYGNEGCGWAGWLLSPRGRSPSGWPTDPRRRSRNLG 120
Db 61 RRQIPKARRPEGRWTAQPGYPWLYGNEGCGWAGWLLSPRGRSPSGWPTDPRRRSRNLG 120

Qy 121 KVDTLTGCFADLMGYIPLVGAPLGGAARALAHGVRVLEGVNYATGNLPGCSFSIFLLA 180
Db 121 KVDTLTGCFADLMGYIPLVGAPLGGAARALAHGVRVLEGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 6
US-08-231-368-36
; Sequence 36, Application US/08231368
; Patent No. 5756312
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/759,575
```

```
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-231-368-36

Query Match 100.0%; Score 1040; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKKKNTNRRPQDVKFGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
Db 1 MSTNPKPQKKKNTNRRPQDVKFGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
Qy 61 RRQIPKARPEGRGRTWAQGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Db 61 RRQIPKARPEGRGRTWAQGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 7
US-08-440-210-36
; Sequence 36, Application US/08440210
; Patent No. 5766845
; GENERAL INFORMATION:
; APPLICANT: Weiner, Amy J.
; APPLICANT: Houghton, Michael
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,210
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/231,368
; FILING DATE:
; APPLICATION NUMBER: US 07/759,575
; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:

; FILING DATE: 13-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McClung, Barbara G.
; REGISTRATION NUMBER: 33,113
; REFERENCE/DOCKET NUMBER: 0205.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2708
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-231-368-36

Query Match 100.0%; Score 1040; DB 1; Length 3011;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKKKNTNRRPQDVKFGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
Db 1 MSTNPKPQKKKNTNRRPQDVKFGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
Qy 61 RRQIPKARPEGRGRTWAQGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Db 61 RRQIPKARPEGRGRTWAQGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 8
US-09-388-874-2
; Sequence 2, Application US/09388874
; Patent No. 6284249
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; FILE REFERENCE: PMCF97-03A
; CURRENT APPLICATION NUMBER: US/09/388,874
; CURRENT FILING DATE: 1999-09-02
; EARLIER APPLICATION NUMBER: PCT/FR98/00448
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 97/02,887
; EARLIER FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Virus
; US-09-388-874-2

Query Match 100.0%; Score 1040; DB 3; Length 3011;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPQKKKNTNRRPQDVKFGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
Db 1 MSTNPKPQKKKNTNRRPQDVKFGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
Qy 61 RRQIPKARPEGRGRTWAQGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Db 61 RRQIPKARPEGRGRTWAQGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191
```

Db 181 LLSCLTVPASA 191

RESULT 9  
US-09-046-604-36  
; Sequence 36, Application US/0904604  
; Patent No. 6303292  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, Amy J.  
; APPLICANT: Houghton, Michael  
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/046,604  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,368  
; FILING DATE:  
; APPLICATION NUMBER: US 07/759,575  
; FILING DATE: 13-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0205.001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2708  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3011 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-046-604-36

Query Match 100.0%; Score 1040; DB 3; Length 3011;  
Best Local Similarity 100.0%; Pred. No. 1.3e-92;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNTNRRPQDVKPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60  
Db 1 MSTNPKQKKNTNRRPQDVKPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60

Qy 61 RRQPIPKARPEGRTPAQPGYPWPPLYGNECGWAGLLSPRGSRPSWGPTDPRRRSRNLG 120  
Db 61 RRQPIPKARPEGRTPAQPGYPWPPLYGNECGWAGLLSPRGSRPSWGPTDPRRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

RESULT 11  
US-08-850-328-1  
; Sequence 1, Application US/08850328  
; Patent No. 6379886  
; GENERAL INFORMATION:  
; APPLICANT: TAKAHAMA, Y.  
; APPLICANT: SHIRAIISHI, J.  
; TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS  
; TITLE OF INVENTION: C VIRUS INFECTION  
; NUMBER OF SEQUENCES: 13

Db 181 LLSCLTVPASA 191

RESULT 10  
US-08-440-549-10  
; Sequence 10, Application US/08440549

```
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & FOERSTER
;; STREET: 2000 Pennsylvania Avenue, NW
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20006-1888
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: Windows
;; SOFTWARE: FastSeq for Windows Version 2.0b
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/850,328
;; FILING DATE: 02-MAY-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mays, Thomas D
;; REGISTRATION NUMBER: 34,524
;; REFERENCE/DOCKET NUMBER: 32273-20004.00
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-887-1500
;; TELEFAX: 202-822-0168
;; TELEX: 90-4030
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3011 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-850-328-1

Query Match 100.0%; Score 1040; DB 3; Length 3011;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPKKKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
DB 1 MSTNPKPKKKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
DB 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSLCLTVPASA 191
DB 181 LLSLCLTVPASA 191

RESULT 13
US-07-681-703B-12
; Sequence 12, Application US/07681703B
; Patent No. 5443965
; GENERAL INFORMATION:
; APPLICANT: Reyes, Gregory
; APPLICANT: Kim, Jungsu P.
; APPLICANT: Moekli, Randolph
; TITLE OF INVENTION: Hepatitis C Virus Epitopes
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Ave., Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/681,703B
; FILING DATE: 05-APR-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 505,611
; FILING DATE: 06-APR-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 594,854
; FILING DATE: 09-OCT-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 4600-076.21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-681-703B-12

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: MORRISON & FOERSTER
;; STREET: 2000 Pennsylvania Avenue, NW
;; CITY: Washington
;; STATE: DC
;; COUNTRY: USA
;; ZIP: 20006-1888
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: Windows
;; SOFTWARE: FastSeq for Windows Version 2.0b
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/850,328
;; FILING DATE: 02-MAY-1997
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Mays, Thomas D
;; REGISTRATION NUMBER: 34,524
;; REFERENCE/DOCKET NUMBER: 32273-20004.00
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202-887-1500
;; TELEFAX: 202-822-0168
;; TELEX: 90-4030
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 3011 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; US-08-850-328-1

Query Match 100.0%; Score 1040; DB 3; Length 3011;
Best Local Similarity 100.0%; Pred. No. 1.3e-92;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKPKKKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
DB 1 MSTNPKPKKKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
DB 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSLCLTVPASA 191
DB 181 LLSLCLTVPASA 191

RESULT 12
US-09-916-359-2
; Sequence 2, Application US/09916359
; Patent No. 6538123
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; FILE REFERENCE: PMCF97-03A
; CURRENT APPLICATION NUMBER: US/09/916,359
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 09/388,874
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: 97/02,887
; PRIOR FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
```

Query Match 99.6%; Score 1036; DB 1; Length 215;  
Best Local Similarity 99.5%; Pred. No. 1.3e-93;  
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPDVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60  
Db 1 MGTNPKQKKNKNTNRRPDVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60

Qy 61 RRQPIPKARRPEGRWTAQPGYPMPLYGNECGWAGWLLSPRGRPSWGPTDPRRSNGLG 120  
Db 61 RRQPIPKARRPEGRWTAQPGYPMPLYGNECGWAGWLLSPRGRPSWGPTDPRRSNGLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYYATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

## RESULT 14

US-07-681-703B-14

; Sequence 14, Application US/07681703B

; Patent No. 5443965

; GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory

; APPLICANT: Kim, Jungsuh P.

; APPLICANT: Moeckli, Randolph

; TITLE OF INVENTION: Hepatitis C Virus Epitopes

; NUMBER OF SEQUENCES: 55

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger &amp; Associates

; STREET: 350 Cambridge Ave., Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/681.703B

; FILING DATE: 05-APR-1991

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 505,611

; FILING DATE: 06-APR-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 594,854

; FILING DATE: 09-OCT-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Fabian, Gary R.

; REGISTRATION NUMBER: 33,875

; REFERENCE/DOCKET NUMBER: 4600-076.21

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 324-0880

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 215 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-681-703B-14

Query Match

Best Local Similarity 99.6%; Score 1036; DB 1; Length 215;

Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPDVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60

Db 1 MGTNPKQKKNKNTNRRPDVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60

Qy 61 RRQPIPKARRPEGRWTAQPGYPMPLYGNECGWAGWLLSPRGRPSWGPTDPRRSNGLG 120

Db 61 RRQPIPKARRPEGRWTAQPGYPMPLYGNECGWAGWLLSPRGRPSWGPTDPRRSNGLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYYATGNLPGCSFSIFLLA 180

Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191

Db 181 LLSCLTVPASA 191

## RESULT 15

PCT-US91-02370-12

; Sequence 12, Application PC/TUS9102370

; GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory

; APPLICANT: Kim, Jungsuh P.

; APPLICANT: Moeckli, Randolph

; APPLICANT: Simonsen, Christian C.

; TITLE OF INVENTION: Hepatitis C Virus Epitopes

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Peter J. Dehlinger

; STREET: P.O. BOX 60850

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US91/02370

; FILING DATE: 19910405

; CLASSIFICATION: 435.5

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 505,611

; FILING DATE: 06-APR-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 594,854

; FILING DATE: 09-OCT-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Fabian, Gary R.

; REGISTRATION NUMBER: 33,875

; REFERENCE/DOCKET NUMBER: 4600-076.41

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 323-8302

; INFORMATION FOR SEQ ID NO: 12:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 215 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

PCT-US91-02370-12

Query Match 99.6%; Score 1036; DB 5; Length 215;

Best Local Similarity 99.5%; Pred. No. 1.3e-93;

Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPDVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60

Query Match

Best Local Similarity 99.5%; Pred. No. 1.3e-93;

Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPDVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60

Qy	121	KVIDTLTCGFADLMGYIPLV	GADPLGGAARALAHGVRVLE	DGVNYATGNLPGCSFSIFLLA	180
Db	121	KVIDTLTCGFADLMGYIPLV	GADPLGGAARALAHGVRVLE	DGVNYATGNLPGCSFSIFLLA	180
Qy	181	LLSCLTVPASA	191		
Db	181	LLSCLTVPASA	191		

Search completed: August 19, 2005, 12:44:53  
Job time : 21 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 12:28:45 ; Search time 117.5 Seconds  
(without alignments)  
628.691 Million cell updates/sec

Title: US-10-770-117-4

Perfect score: 1040

Sequence: 1 MSTNPKPQKGNKNTNRPO.....CSFSIFLLALLSLCTVPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040	100.0	191	3	AAY82999 Hepatitis
2	1040	100.0	2816	2	AAR34009 HCV-1 pol
3	1040	100.0	2955	2	AAY14975 Amino aci
4	1040	100.0	2955	3	AAB18541 Polypeptide
5	1040	100.0	3011	2	AAR21519 Compiled
6	1040	100.0	3011	2	AAR90931 Hepatitis
7	1040	100.0	3011	2	AAR34480 HCV poly
8	1040	100.0	3011	2	AAR40038 HCV poly
9	1040	100.0	3011	5	AAR22049 Hepatitis
10	1040	100.0	3011	8	ADL23107 Hepatitis
11	1040	100.0	3011	8	ADR29357 Hepatitis
12	1036	99.6	190	7	ADP39661 Hepatitis
13	1036	99.6	215	2	AAR14557 HCV Capsi
14	1036	99.6	215	2	AAR14558 Polypepti
15	1036	99.6	215	2	AAR81945 Hepatitis
16	1036	99.6	215	2	AAR81944 Hepatitis
17	1036	99.6	217	2	AAR70828 Hepatitis
18	1036	99.6	217	2	AAR70829 Insert of
19	1036	99.6	217	2	AAR80517 Insert of
20	1036	99.6	217	2	AAR80516 Hepatitis
21	1035	99.5	190	5	AAB71256 HCV type
22	1035	99.5	191	3	AAY82997 Hepatitis
23	1035	99.5	2955	8	ADN35978 HCV cDNA
24	1031	99.1	191	2	AAR44010 Hepatitis
25	1031	99.1	191	2	AAR92938 Hepatitis

ALIGNMENTS

RESULT 1

AAY82999  
ID AAY82999 standard; peptide; 191 AA.

XX AAY82999;

XX 12-SEP-2003 (revised)

DT 04-JUL-2000 (first entry)

DE Hepatitis C virus core polypeptide.

XX Hepatitis C virus; HCV; core polypeptide; antigen; epitope; agretope;

KW anchor residue; cytotoxic T lymphocyte; CTL; immune response; MHC;

KW Major histocompatibility complex; vaccine; treatment.

XX Hepatitis C virus; (HCV).

XX WO200011186-A1.

XX 02-MAR-2000.

PF 17-AUG-1999; 99WO-US018674.

PR 21-AUG-1998; 98US-0097446P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Berzofsky JA, Sarobe P, Pendleton CD, Feinstein SM, Arichi T;

PI Major ME;

XX WPI; 2000-246569/21.

PT Hepatitis C virus polypeptides is useful as a vaccine for treating

PT Hepatitis C virus infection and for activating cytotoxic T lymphocytes.

XX Disclosure; Page 72-73; 78pp; English.

XX Peptide fragments of the Hepatitis C virus (HCV) core polypeptide can be used to elicit an immune response. A cytotoxic T-lymphocyte (CTL) response is present in patients acutely and chronically infected by HCV. CTL's recognise viral antigens as peptides presented by Class I molecules of the major histocompatibility complex (MHC). These peptide antigens are typically 8-10 amino acids long. These peptide antigens typically have a sequence pattern or motif depending upon which particular MHC antigen they are bound and presented by. The motifs are based on the presence in precise positions in the peptide sequence of several amino acids (agretopic residues) called anchor residues, responsible for interactions between peptide and MHC molecule, as well as other secondary position

CC that may help to stabilise the interaction. Thus, single amino acid  
 CC changes within the peptide sequence can enhance the immune response.  
 CC Peptides derived from the HCV core polypeptide having substituted amino  
 CC acids within their sequence, specifically a hepatitis C virus core  
 CC polypeptide comprising an L to an A substitution at amino acid position  
 CC 139 enhance the immune response against HCV. They can therefore be used  
 CC as a vaccine or for treating HCV infection. (Updated on 12-SEP-2003 to  
 CC standardise OS field)

XX  
 SQ Sequence 191 AA;  
 Query Match 100.0%; Score 1040; DB 3; Length 191;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-90;  
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPDVKFPGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
 DB 1 MSTNPKQKKNKNTNRRPDVKFPGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
 QY 61 RRQPIKARPEGRGRTWAQPGYPMPLYGNECGWAGWLLSPRGSRSRPSWGPTDPRRSRNLG 120  
 DB 61 RRQPIKARPEGRGRTWAQPGYPMPLYGNECGWAGWLLSPRGSRSRPSWGPTDPRRSRNLG 120  
 QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTVPASA 191  
 DB 181 LLSCLTVPASA 191

RESULT 2  
 AAR34009  
 ID AAR34009 standard; protein; 2816 AA.  
 AC AAR34009;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 26-JUL-1993 (first entry)  
 XX  
 XX HCV-1 polypeptide.  
 DE  
 XX Polymerase chain reaction; PCR; amplify; primer; hepatitis C virus; HCV;  
 KW asymptomatic; chronically infected; epitope; viral isolate; domain;  
 KW immunological; cross-reactive; envelope protein; vaccine;  
 KW gp53 (BVDV)/gp55; hog cholera virus; pestivirus; NS1; flavivirus.  
 XX  
 OS Hepatitis C virus.  
 XX  
 PN WO9306126-A1.  
 XX  
 PD 01-APR-1993.  
 XX  
 PF 11-SEP-1992; 92WO-US007683.  
 XX  
 PR 13-SEP-1991; 91US-00759575.  
 XX  
 XX (CHIR ) CHIRON CORP.  
 PA  
 XX Weiner AJ, Houghton M;  
 PI  
 XX WPI; 1993-117468/14.  
 DR  
 XX Immuno-reactive hepatitis C virus polypeptide compsns. - contg. at least  
 PT 2 sequences from the first variable domain of distinct HCV isolates.  
 XX  
 PS Disclosure; Fig 9; 106pp; English.  
 XX  
 XX This sequence represents the entire hepatitis C virus polypeptide. HCV is  
 CC a member of the flavivirus family and appears to encode a basic  
 CC polypeptide domain ("C") at the N-terminal of the viral polypeptide,  
 CC followed by two glycoprotein domains ("E1", "E2/NS1"), upstream of the

CC nonstructural genes NS2 through NS5. See also AAQ39134-48, AAR33982- 4008  
 CC and AAR38088-89. (Updated on 25-MAR-2003 to correct PN field.)

XX  
 SQ Sequence 2816 AA;  
 Query Match 100.0%; Score 1040; DB 2; Length 2816;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-89;  
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPDVKFPGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
 DB 1 MSTNPKQKKNKNTNRRPDVKFPGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
 QY 61 RRQPIKARPEGRGRTWAQPGYPMPLYGNECGWAGWLLSPRGSRSRPSWGPTDPRRSRNLG 120  
 DB 61 RRQPIKARPEGRGRTWAQPGYPMPLYGNECGWAGWLLSPRGSRSRPSWGPTDPRRSRNLG 120  
 QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTVPASA 191  
 DB 181 LLSCLTVPASA 191

RESULT 3  
 AAY14975  
 ID AAY14975 standard; protein; 2955 AA.  
 AC AAY14975;  
 XX  
 XX 20-MAR-2003 (revised)  
 DT 08-NOV-1999 (first entry)  
 XX  
 XX Amino acid sequence of HCV-1 ORF.  
 DE  
 XX Hepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH;  
 KW HCV infection; vaccine.  
 XX  
 OS Hepatitis C virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 441 /note= "encoded by TT"  
 FT Misc-difference 461 /note= "encoded by CCCC"  
 XX  
 PN EP939128-A2.  
 XX  
 PD 01-SEP-1999.  
 XX  
 PF 17-SEP-1990; 99EP-00101746.  
 XX  
 PR 15-SEP-1989; 89US-00408045.  
 PR 21-DEC-1989; 89US-00456142.  
 PR 17-SEP-1990; 90EP-00310149.  
 XX  
 XX (OYAA/) OYA A.  
 PA (CHIR ) CHIRON CORP.  
 PA  
 XX Miyamura T, Saito I, Houghton M, Weiner AJ, Han J, Kolberg JA;  
 PI Cha T, Irvine BD;  
 XX  
 DR WPI; 1999-480843/41.  
 DR N-PSDB; AA207656.  
 XX  
 PT New Hepatitis C Virus isolates, useful for diagnosis of hepatitis  
 PT infections and development of vaccines.  
 XX  
 PS Disclosure; Fig 12; 132pp; English.  
 XX  
 XX The invention provides two new isolates of hepatitis C virus (HCV), J1

CC and J7. These two isolates comprise nucleotide and amino acid sequences  
CC that are distinct from the HCV isolate HCV-1. The nucleotide sequences  
CC may be used to detect non-A, non-B HCV (NANBH) polynucleotides by  
CC hybridisation for diagnosis of NANBH infections. They may also be used to  
CC screen blood donors, donated blood and blood products for this infection.  
CC The isolates may also be used to isolate other naturally occurring  
CC variants of the virus. The polypeptides may be used as a vaccine for  
CC administration to patients to protect against infection with NANBH. The  
CC present sequence represents the amino acid sequence of HCV-1 ORF.  
CC (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to  
CC correct PR field.)  
XX  
XX Sequence 2955 AA;

Query Match 100.0%; Score 1040; DB 2; Length 2955;  
Best Local Similarity 100.0%; Pred. No. 3.3e-89;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSTNPKPKQKKNKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
DB 1 MSTNPKPKQKKNKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
QY 61 RRQPIPKARRPEGRWTQAPGYPMPLYGNEGCGWAGWLLSPRGRPSWGPTDPRRSRLG 120  
DB 61 RRQPIPKARRPEGRWTQAPGYPMPLYGNEGCGWAGWLLSPRGRPSWGPTDPRRSRLG 120  
QY 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGDGVNATGNLPGCSFSIFLLA 180  
DB 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGDGVNATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTVPASA 191  
DB 181 LLSCLTVPASA 191

RESULT 4  
AAB18541  
ID AAB18541 standard; protein; 2955 AA.  
XX AC AAB18541;  
XX  
XX 15-JAN-2001 (first entry)  
XX Polypeptide encoded by sense strand of HCV.  
XX Hepatitis C virus; HCV; antisense polynucleotide; polypeptide;  
XX viral infectivity; viral replication.  
XX Hepatitis C virus.  
XX EPI034785-A2.  
XX  
XX 13-SEP-2000.  
XX 16-MAR-1990; 2000EP-00109602.  
XX 17-MAR-1989; 89US-00325338.  
XX 20-APR-1989; 89US-003411334.  
XX 18-MAY-1989; 89US-00355002.  
XX 16-MAR-1990; 90EP-00302866.  
XX (CHIR ) CHIRON CORP.  
XX Houghton M, Choo Q, Kuo G;  
XX WPI; 2000-566891/53.  
XX N-PSDB; AAA75297.  
XX Novel composition comprising a hepatitis C virus antisense polynucleotide  
XX which is complementary to or corresponds to a sense strand of the virus  
XX genome, and selectively hybridizes to it.  
XX Example; Fig 17; 75pp; English.

XX The specification describes a pharmaceutical composition which comprises  
CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is  
CC characterized by a positive stranded RNA genome which has 40% homology at  
CC the polypeptide level to a HCV polypeptide. The antisense polynucleotide  
CC binds to cellular polynucleotides which enhance and/or are required for  
CC viral infectivity, replicative ability or chronicity. The antisense  
CC polynucleotides may also be designed to bind with high specificity, to be  
CC of increased stability, to be stable and to have low toxicity. The  
CC composition also comprises an agent which causes viral RNA to be  
CC inactive. The composition is used for preventing HCV replication in a  
CC system. The present sequence is encoded by a novel HCV CDNA sequence,  
CC which is used in the course of the invention  
XX  
XX Sequence 2955 AA;

Query Match 100.0%; Score 1040; DB 3; Length 2955;  
Best Local Similarity 100.0%; Pred. No. 3.3e-89;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSTNPKPKQKKNKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
DB 1 MSTNPKPKQKKNKNTNRRPQDVKFGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
QY 61 RRQPIPKARRPEGRWTQAPGYPMPLYGNEGCGWAGWLLSPRGRPSWGPTDPRRSRLG 120  
DB 61 RRQPIPKARRPEGRWTQAPGYPMPLYGNEGCGWAGWLLSPRGRPSWGPTDPRRSRLG 120  
QY 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGDGVNATGNLPGCSFSIFLLA 180  
DB 121 KVDTLTTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGDGVNATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTVPASA 191  
DB 181 LLSCLTVPASA 191

RESULT 5  
AAR21519  
ID AAR21519 standard; protein; 3011 AA.  
XX AC AAR21519;  
XX  
XX 24-OCT-2003 (revised)  
XX 25-MAR-2003 (revised)  
XX 22-JUN-1992 (first entry)  
XX Compiled HCV sequence.  
XX HCV1; serum; gtl1.  
XX  
XX Hepatitis C virus type 1.  
XX  
XX Key Location/Qualifiers  
XX Misc-difference 9 /label= ARG  
XX Misc-difference 11 /label= THR  
XX Misc-difference 176 /label= THR  
XX Misc-difference 334 /label= VAL  
XX Misc-difference 603 /label= ILE  
XX Misc-difference 848 /label= (ASN)  
XX Misc-difference 1114 /label= SER  
XX Misc-difference 1117 /label= THR  
XX Misc-difference 1276 /label= LEU  
XX Misc-difference 1328

```

FT      /label= (VAL)
FT Misc-difference 1454
FT      /label= TYR
FT Misc-difference 1471
FT      /label= (SER)
FT Misc-difference 1877
FT      /label= (GLY)
FT Misc-difference 1948
FT      /label= (HIS)
FT Misc-difference 1949
FT      /label= (CYS)
FT Misc-difference 2021
FT      /label= (VAL)
FT Misc-difference 2349
FT      /label= (SER)
FT Misc-difference 2385
FT      /label= (PHE)
FT Misc-difference 2386
FT      /label= (ALA)
FT Misc-difference 2502
FT      /label= (PHE)
FT Misc-difference 2690
FT      /label= (GLY)
FT Misc-difference 2996
FT      /label= (PRO)
XX
XX WO9202642-A.
PN
XX
XX
XX 20-FEB-1992.
XX
XX 10-AUG-1990; 90US-00566209.
XX
XX 10-AUG-1990; 90US-00566209.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Houghton M, Choo Q, Kuo G, Weiner AJ, Ureda MS, Irvine BD;
XX Kolberg JA;
XX
XX WPI; 1992-080094/10.
XX N-PSDB; AAQ21744.
XX
XX Reagents for isolating, amplifying and detecting HCV polynucleotide(s) -
XX used to monitor spread of blood-borne non-a, non-b hepatitis virus
XX infection and screen blood samples for virus.
XX
XX Disclosure; Fig 1; 67pp; English.
XX
XX Heterogeneities in cloned DNAs of HCV1 are indicated by the amino acid
XX indicated in the features, the parentheses indicated that the
XX heterogeneity was detected at or near to the 5'- or 3'-end of the HCV1, a
XX the clone. The sequence is derived from a composite HCV cDNA from HCV1, a
XX prototypic HCV. The DNA sequence is based upon sequence information
XX derived from a no. of HCV cDNA clones, which were isolated from a no. of
XX HCV cDNA libraries, including the "c" library present in lambda gt11
XX (ATCC No.40394), and from human serum. The HCV cDNA clones were isolated
XX by methods described in WO9014436. The clones from which the sequence was
XX derived are 5'clone32, b114a, 18g, ag30a, CA205a, CA290a, CA216a, p14a,
XX CA167b, CA156e, CA84a, CA59a, K9-1 (also called K9-1), 26j, 13i, 12f,
XX 14i, 11b, 7f, 8h, 33c, 40b, 37b, 35, 36, 81, 32, 33b, 25c, 14c, 8f,
XX 33f, 33g, 39c, 35f, 19g, 26g, 15e, b5a, 16ih, 6k and p131jh. (Updated on
XX 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI
XX field.) (Updated on 24-OCT-2003 to standardise OS field)
XX
XX Sequence 3011 AA;
XX
XX Query Match 100.0%; Score 1040; DB 2; Length 3011;
XX Best Local Similarity 100.0%; Pred. No. 3.4e-89;
XX Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MSTNPKQKKNRNTNRPPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
XX
XX 1 MSTNPKQKKNRNTNRPPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

```

```

QY 61 RRQIPKARRPRGRTWAOPGYDPLVGNEGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
Db |||||
QY 61 RRQIPKARRPRGRTWAOPGYDPLVGNEGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120
Db |||||
QY 121 KVIDTILTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db |||||
QY 121 KVIDTILTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db |||||
QY 181 LLSCLTVPASA 191
Db |||||
Db 181 LLSCLTVPASA 191

RESULT 6
AAR90931
ID AAR90931 standard; protein; 3011 AA.
XX
XX AAR90931;
XX
XX 25-MAR-2003 (revised)
DT 15-MAY-1996 (first entry)
XX
XX Hepatitis C virus polyprotein.
XX
XX Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;
XX antibodies.
XX
XX Hepatitis C virus.
XX
XX Key Location/Qualifiers
FT Misc-difference 1..122 /label= antigen
FT /note= "C22; AAR90936"
FT Misc-difference 199..328
FT /label= antigen
FT /note= "S2; AAR90935"
FT Misc-difference 1192..1457
FT /label= antigen
FT /note= "C33c; AAR90932"
FT Misc-difference 1569..1931
FT /label= antigen
FT /note= "C100; AAR90933"
FT Misc-difference 2054..2464
FT /label= antigen
FT /note= "NS5; AAR90934"
XX
XX EP693687-A1.
XX
XX 24-JAN-1996.
XX
XX 03-APR-1991; 95EP-00114016.
XX
XX 04-APR-1990; 90US-00504352.
XX (CHIR ) CHIRON CORP.
XX
XX Houghton M, Choo Q, Kuo G;
XX WPI; 1996-117956/13.
XX N-PSDB; AAT12710.
XX
XX Combinations of synthetic Hepatitis C Virus antigens - provide more
XX effective diagnosis of Non-A, Non-B Hepatitis.
XX
XX Disclosure; Fig 1(A-Y); 53pp; English.
XX
XX The combination comprises an HCV antigen from the C domain (pref. C22 -
XX AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -
XX AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90935) or NS5
XX (AAR90934) domain. The antigens may in the form of a fusion protein, a
XX simple physical mixture, or the individual antigens commonly bound to a
XX solid matrix. They are pref. prepd. by recombinant DNA techniques

```

CC	(primers are given in AAT12711-T12716), but can be synthesised or									
CC	isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003									
CC	to correct PF field.)									
XX										
SQ	Sequence 3011 AA;									
	Query Match	100.0%;	Score 1040;	DB 2;	Length 3011;					
	Best Local Similarity	100.0%;	Pred. No. 3.4e-89;							
	Matches 191;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;					
QY	1	MSTNPKPKKKNTNRRPDDVKFPGGGQIVGGVYLLPRRGPRGLGVRATRTKTSERSOPRG	60							
DB	1	MSTNPKPKKKNTNRRPDDVKFPGGGQIVGGVYLLPRRGPRGLGVRATRTKTSERSOPRG	60							
QY	61	RRQIPKARRPEGRTWAQPGYPWLYNGEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120							
DB	61	RRQIPKARRPEGRTWAQPGYPWLYNGEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120							
QY	121	KVIDTLTCGFADLMGYIPLVCAPIGGARALAHGVRVLEEDGVNATGNLPGCSFSIFLLA	180							
DB	121	KVIDTLTCGFADLMGYIPLVCAPIGGARALAHGVRVLEEDGVNATGNLPGCSFSIFLLA	180							
QY	181	LLSCLTVPASA 191								
DB	181	LLSCLTVPASA 191								
RESULT 7										
AAW34480										
ID	AAW34480 standard; protein; 3011 AA.									
XX	AAW34480;									
XX										
DT	25-MAR-2003 (revised)									
DT	16-MAR-1998 (first entry)									
XX										
DE	HCV polyprotein.									
XX										
KW	PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;									
KW	C domain; S domain; NS5; HCV polyprotein; anti-HCV antibody; detection;									
KW	NS4.									
XX										
OS	Hepatitis C virus.									
XX										
Key	Location/Qualifiers									
FT	Misc-difference	366	/note=	"can optionally be Arg"						
FT	Misc-difference	372	/note=	"can optionally be Thr"						
FT	Misc-difference	867	/note=	"can optionally be Thr"						
FT	Misc-difference	1341	/note=	"can optionally be Val"						
FT	Misc-difference	2148	/note=	"can optionally be Ile"						
FT	Misc-difference	2883	/note=	"can optionally be Asn"						
FT	Misc-difference	3681	/note=	"can optionally be Ser"						
FT	Misc-difference	3690	/note=	"can optionally be Thr"						
FT	Misc-difference	4167	/note=	"can optionally be Leu"						
FT	Misc-difference	4323	/note=	"can optionally be Val"						
FT	Misc-difference	4701	/note=	"can optionally be Tyr"						
FT	Misc-difference	4752	/note=	"can optionally be Ser"						
FT	Misc-difference	5970	/note=	"can optionally be Gly"						
FT	Misc-difference	6183	/note=	"can optionally be His"						

Db	1	MSTNPKQKKNKNTNRRPDQVKFGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG	60
QY	61	RRQPIPKARRPEGRTHAQQGYFWPLVYNGEGCGWAGWLLSPRGSRPDSWGPTDPRRRSRNLG	120
Db	61	RRQPIPKARRPEGRTHAQQGYFWPLVYNGEGCGWAGWLLSPRGSRPDSWGPTDPRRRSRNLG	120
QY	121	KVIDILTCGFADLMGYIPIVGPAPLGGAARALAHGVVRVLEDDGVNYATGNLPGCSFSIFLLA	180
Db	121	KVIDILTCGFADLMGYIPIVGPAPLGGAARALAHGVVRVLEDDGVNYATGNLPGCSFSIFLLA	180
QY	181	LLSCLTVTPASA 191	
Db	181	LLSCLTVTPASA 191	
RESULT 8			
AAW40038			
ID	AAW40038 standard; protein; 3011 AA.		
XX			
AC	AAW40038;		
XX			
DT	26-MAY-1998 (first entry)		
XX			
DE	HCV polypeptide.		
XX			
KW	Hepatitis C virus C domain; HCV; C antigen; immunological activity;		
KW	NS3 domain; NS4 domain; S domain; NS5 domain.		
XX			
OS	Hepatitis C virus.		
XX			
FH	Key	Location/Qualifiers	
FT	Domain	1..120	
FT		/label= C_domain	
FT	Modified-site	9	
FT		/note= "As given in the specification this amino acid can	
FT		also be Arg"	
FT	Modified-site	11	
FT		/note= "As given in the specification this amino acid can	
FT		also be Thr"	
FT	Domain	120..400	
FT		/label= S_domain	
FT	Modified-site	174	
FT		/note= "As given in the specification this amino acid can	
FT		also be Thr"	
FT	Modified-site	334	
FT		/note= "As given in the specification this amino acid can	
FT		also be Val"	
FT	Modified-site	603	
FT		/note= "As given in the specification this amino acid can	
FT		also be Ile"	
FT	Modified-site	847	
FT		/note= "As given in the specification this amino acid can	
FT		also be Asn"	
FT	Domain	1050..1640	
FT		/label= NS3_domain	
FT	Modified-site	1114	
FT		/note= "As given in the specification this amino acid can	
FT		also be Ser"	
FT	Modified-site	1217	
FT		/note= "As given in the specification this amino acid can	
FT		also be Thr"	
FT	Modified-site	1276	
FT		/note= "As given in the specification this amino acid can	
FT		also be Leu"	
FT	Modified-site	1328	
FT		/note= "As given in the specification this amino acid can	
FT		also be Val"	
FT	Modified-site	1452	
FT		/note= "As given in the specification this amino acid can	
FT		also be Tyr"	
FT	Modified-site	1472	
FT		/note= "As given in the specification this amino acid can	
FT		also be Ser"	

FT	Domain	1640..2000	
FT	Modified-site	/label= NS4_domain	
FT		1877	
FT		/note= "As given in the specification this amino acid can also be Gly"	
FT	Modified-site	1948	
FT		/note= "As given in the specification this amino acid can also be His"	
FT	Modified-site	1949	
FT		/note= "As given in the specification this amino acid can also be Cys"	
FT	Domain	2000..3011	
FT		/label= NS5_domain	
FT	Modified-site	2021	
FT		/note= "As given in the specification this amino acid can also be Val"	
FT	Modified-site	2348	
FT		/note= "As given in the specification this amino acid can also be Ser"	
FT	Modified-site	2385	
FT		/note= "As given in the specification this amino acid can also be Phe"	
FT	Modified-site	2386	
FT		/note= "As given in the specification this amino acid can also be Ala"	
FT	Modified-site	2502	
FT		/note= "As given in the specification this amino acid can also be Phe"	
FT	Modified-site	2690	
FT		/note= "As given in the specification this amino acid can also be Gly"	
FT	Modified-site	2921	
FT		/note= "As given in the specification this amino acid can also be Gly"	
FT	Modified-site	2996	
FT		/note= "As given in the specification this amino acid can also be Pro"	
XX			
FN	US5712087-A.		
XX			
XX	27-JAN-1998.		
XX			
DF	12-MAY-1995;	95US-00440519.	
XX			
XX	04-APR-1990;	90US-00504352.	
PR	07-JUL-1992;	92US-00910760.	
XX			
PA	(CHIR ) CHIRON CORP.		
XX			
FI	Kuo G, Houghton M, Choo Q;		
XX			
XX	WPI; 1998-119973/11.		
DR	N-PSDB; AAV09989.		
DR			
XX			
PT	Immunoassays for hepatitis C virus antibodies - using combinations of antigenic fragments of HCV polypeptide.		
FT			
XX	Disclosure; Fig 1; 59pp; English.		
XX			
CC	This sequence represents the hepatitis C virus (HCV) polypeptide which is used in the construction of novel combinations of HCV antigens that have a broader range of immunological activity than any single HCV antigen. An example of such an antigen given in this specification comprises a first antigen containing at least 8 amino acids of the C domain of the HCV polypeptide and a second antigen comprising at least 8 amino acids of the NS3 domain, the NS4 domain, the S domain or the NS5 domain of the HCV polypeptide in the form of a fusion protein, a physical mixture or bound to a solid matrix. Note: The features given in the specification as represented in the feature table of AAM40038 differ from the positions indicated in Figure 1		
CC			
XX	Sequence 3011 AA;		
SQ			

Query Match 100.0%; Score 1040; DB 2; Length 3011;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-89;  
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
 DB 1 MSTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

QY 61 RROPIPKARRPEGTWAQPGYPMPLYNCGCGWAGWLLSPRGSPPSGWPTDPRRRSRNLG 120  
 DB 61 RROPIPKARRPEGTWAQPGYPMPLYNCGCGWAGWLLSPRGSPPSGWPTDPRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191  
 DB 181 LLSCLTVPASA 191

RESULT 9  
 AAE22049  
 ID AAE22049 standard; protein; 3011 AA.  
 XX  
 AC AAE22049;  
 DT 16-JUL-2002 (first entry)  
 DE Hepatitis C virus (HCV) polyprotein.  
 KW Hepatitis C virus; HCV; antigen; C domain; polyprotein; NS3 domain;  
 KW NS4 domain; S domain; NS5 domain.  
 OS Hepatitis C virus.

Key Location/Qualifiers  
 PH Domain 1..122  
 FT Domain /label= C\_domain  
 FT Domain 199..328  
 FT Region /label= S\_domain  
 FT Region 1192..1931  
 FT Region /note= "c200 polypeptide"  
 FT Region 1192..1457  
 FT Region /note= "NS3 domain antigen"  
 FT Region 1569..1931  
 FT Region /note= "NS4 antigen"  
 FT Region 2054..2464  
 FT Region /note= "NS5 antigen"

US6312889-B1.  
 XX  
 PN 06-NOV-2001.  
 XX  
 PD 12-MAY-1995; 95US-00440549.  
 XX  
 PF 04-APR-1990; 90US-00504352.  
 XX  
 PR 07-JUL-1992; 92US-00910760.  
 XX  
 XX (CHIR ) CHIRON CORP.  
 XX  
 XX Houghton M, Choo Q, Kuo G;  
 FI WPI; 2002-040268/05.  
 XX  
 DR N-PSDB; AAD35043.  
 XX  
 XX Combination of hepatitis C viral (HCV) antigens, useful in improved  
 PT immunoassay for detecting HCV antibodies.  
 XX  
 XX Example 1; Col 45-60; 58pp; English.  
 PS  
 XX The invention relates to combination of hepatitis C viral (HCV) antigens  
 CC that have a broader range of immunological reactivity than any single HCV

CC antigen. The combinations consist of an antigen from the C domain of the  
 CC HCV polyprotein, and at least one additional HCV antigen from either the  
 CC NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in  
 CC the form of fusion protein, a simple physical mixture, or the individual  
 CC antigens commonly bound to a solid matrix. The combinations of antigens  
 CC provides broad range immunoassays for anti-HCV antibodies. The invention  
 CC therefore provides a method for detecting antibodies to HCV in a mammal  
 CC suspected of containing such antibodies. The present sequence is HCV  
 CC polyprotein. Note: This sequence SEQ.ID.NO:10 is stated to be similar to  
 CC the sequence shown in Fig 1 (AAE22052) of the specification. However  
 CC these sequences differ  
 XX

XX Sequence 3011 AA;  
 SQ Query Match 100.0%; Score 1040; DB 5; Length 3011;  
 Best Local Similarity 100.0%; Pred. No. 3.4e-89;  
 Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
 DB 1 MSTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

QY 61 RROPIPKARRPEGTWAQPGYPMPLYNCGCGWAGWLLSPRGSPPSGWPTDPRRRSRNLG 120  
 DB 61 RROPIPKARRPEGTWAQPGYPMPLYNCGCGWAGWLLSPRGSPPSGWPTDPRRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191  
 DB 181 LLSCLTVPASA 191

RESULT 10  
 ADL23107  
 ID ADL23107 standard; protein; 3011 AA.  
 XX  
 AC ADL23107;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Hepatitis C virus protein sequence SeqID 2.  
 XX  
 KW hepatitis C virus; HCV; viral envelope; E2 glycoprotein;  
 KW low density lipoprotein; LDL; HCV infection; hypercholesterolaemia;  
 KW hyperlipidaemia; coronary heart disease; hepatotropic; virucidal;  
 KW plasma lipoprotein.  
 XX  
 OS Hepatitis C virus.  
 XX  
 PN WO2004003141-A2.  
 XX  
 XX 08-JAN-2004.  
 XX  
 XX 24-JUN-2003; 2003WO-US019834.  
 XX  
 XX 28-JUN-2002; 2002US-0392158P.  
 PR  
 XX 27-MAY-2003; 2003US-00445724.  
 XX  
 XX (IOWA ) UNIV IOWA RES FOUND.  
 XX  
 XX Stapleton JT, Wuenschmann S;  
 FI WPI; 2004-083029/08.  
 XX  
 DR N-PSDB; ADL23106.  
 XX  
 XX Use of Hepatitis C Virus E2 glycoprotein in reducing low density  
 PT lipoprotein levels in a subject or in treating or preventing hepatitis C  
 FT virus infection.  
 XX  
 PS Disclosure; SEQ ID NO 2; 135pp; English.

XX This invention relates to a novel method for identifying inhibitors of  
CC the hepatitis C virus (HCV), and also the use of an HCV viral envelope  
CC protein identified as E2 glycoprotein in reducing low density lipoprotein  
CC (LDL) levels. Specifically, it refers to the knowledge that HCV E2  
CC glycoprotein binds to the lipid moiety of human lipoproteins and uses the  
CC natural LDL receptor to bind to cells, thus identifies a novel route by  
CC which HCV gains entry to the cell providing a novel therapeutic target.  
CC The present invention describes screening methods to identify inhibitors  
CC of HCV infection that comprises administering a candidate substance with the  
CC E2 glycoprotein and a plasma lipoprotein in order to determine a  
CC reduction in binding in the presence of an appropriate inhibitor. As  
CC such, these compositions can be used to reduce hypercholesterolemia and  
CC in turn treat hyperlipidaemia and associated conditions including  
CC coronary heart disease. Accordingly, they exhibit hepatotropic and  
CC virucidal activities. This polypeptide sequence is the HCV protein  
CC sequence of the invention.  
XX  
SQ Sequence 3011 AA;  
  
Query Match 100.0%; Score 1040; DB 8; Length 3011;  
Best Local Similarity 100.0%; Pred. No. 3.4e-89;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MSTNPKPKKKNTNRPPQDVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1 MSTNPKPKKKNTNRPPQDVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60  
  
QY 61 RRQPIKARRPEGRTPWAQPGYPPWPLYGNECGWAGLLSPRGSRPWGPTDPRRSRNLG 120  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
61 RRQPIKARRPEGRTPWAQPGYPPWPLYGNECGWAGLLSPRGSRPWGPTDPRRSRNLG 120  
  
QY 121 KVITDTTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
121 KVITDTTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
  
QY 181 LLSCLTVPASA 191  
DB ||||||||||||  
181 LLSCLTVPASA 191  
  
RESULT 11  
ADR29357  
ID ADR29357 standard; protein; 3011 AA.  
XX  
AC ADR29357;  
XX  
DT 18-NOV-2004 (first entry)  
XX  
DE Hepatitis C virus polyprotein precursor protein SEQ ID NO:1.  
XX  
KW anti-hepatitis C virus; anti-HCV; immune response; anti-HIV; virucide;  
KW anti-inflammatory; hepatotropic; immunostimulant; vaccine; immunotherapy;  
KW anti-HCV vaccine; AIDS; hepatitis C virus; HCV; polyprotein precursor.  
XX  
OS Hepatitis C virus.  
XX  
FH Key Location/Qualifiers  
FT Peptide 2594..2602  
FT /note= "specifically claimed peptide in claims 4 and 5"  
XX  
PN W02004071414-A2.  
XX  
XX 26-AUG-2004.  
XX  
XX 03-FEB-2004; 2004WO-US003044.  
XX  
PR 05-FEB-2003; 2003US-0445438P.  
PR 25-MAR-2003; 2003US-04557870P.  
XX  
XX (GENZ ) GENZYME CORP.  
PA (GENO ) GEN HOSPITAL CORP DBA MASSACHUSETTS GEN.  
XX

PI Nicolette CA, Walker BD;  
XX  
DR WPI; 2004-625761/60.  
DR N-PSDB; ADR29358.  
DR GENBANK; M62321.  
XX  
PT New anti- hepatitis C virus (HCV) (AL9) compound, useful as component of  
PT anti-HCV vaccines and to expand immune effector cell specific for cells  
PT expressing HCV epitope.  
XX  
PS Claim 4; SEQ ID NO 1; 69pp; English.  
XX  
XX The present invention describes anti-hepatitis C virus (HCV) (AL9)  
CC compounds (I) having the amino acid sequences of SEQ ID NO:3, 5 and 7  
CC (ADR29359, ADR29361 and ADR29363). Also described: (1) a peptide (II)  
CC comprising the 3011 amino acid sequence of SEQ ID NO:1 (SI, ADR29357),  
CC where in (SI) positions 2594-2602 are Phe, Ile, Leu, Trp, Pro, Glu, Asn,  
CC Lys, and Val (SEQ ID NO:3), respectively or Phe, Leu, Trp, Glu, Ala,  
CC Trp, Lys, and Val (SEQ ID NO:5) respectively; (2) a polynucleotide (III)  
CC that encodes Phe-Leu-Ile-Trp-Pro-Glu-Asn-Lys-Val; Phe-Leu-Pro-Trp-Gly-Ala  
CC -Trp-Lys-Val, or Ala-Leu-Tyr-Asp-Val-Thr-Lys-Leu (SEQ ID NO:3, 5 or 7);  
CC (3) an antibody (IV) that recognises and binds to (I); (4) an immune  
CC effector cell (V) that has been raised in vivo or in vitro or in the  
CC presence and the expense of an antigen presenting cell that presents (I),  
CC in the context of an MHC molecule; (5) a composition (VI) comprising (I),  
CC of SEQ ID NO:3, 5 or 7 where (I) is individually characterised by an  
CC ability to elicit an immune response against the same native ligand; (6)  
CC a composition (VII) comprising (I) and a carrier; (7) a host cell (VIII)  
CC comprising at least one or more (I) of SEQ ID NO:3, 5 or 7 that are  
CC individually characterised by an ability to elicit an immune response  
CC against the same native ligand; and (8) a composition comprising (VIII)  
CC and a carrier. (I) have anti-HIV, virucide, antiinflammatory  
CC hepatotropic and immunostimulant activities, and can be used in vaccines.  
CC (I) can be used for inducing an immune response in a subject which  
CC involves delivering (I) to the subject. (I) is delivered in the context  
CC of an MHC molecule. The MHC molecules presents (I) on the surface of an  
CC antigen presenting cell. (I) is delivered as a polynucleotide that  
CC encodes it. (IV) is useful for immunotherapy. (V) is useful for adoptive  
CC immunotherapy. (VI) is useful for inducing an immune response in a  
CC subject. (I) is useful as components of anti-HCV vaccines and to expand  
CC immune effector cell that are specific for cells expressing HCV epitope.  
CC (I) is useful for preparation of medicament for diagnosis and treatment  
CC of diseases such as AIDS. The present sequence represents the hepatitis C  
CC virus polyprotein precursor, which is used in the exemplification of the  
CC present invention.  
XX  
SQ Sequence 3011 AA;  
  
Query Match 100.0%; Score 1040; DB 8; Length 3011;  
Best Local Similarity 100.0%; Pred. No. 3.4e-89;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MSTNPKPKKKNTNRPPQDVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
1 MSTNPKPKKKNTNRPPQDVKFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60  
  
QY 61 RRQPIKARRPEGRTPWAQPGYPPWPLYGNECGWAGLLSPRGSRPWGPTDPRRSRNLG 120  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
61 RRQPIKARRPEGRTPWAQPGYPPWPLYGNECGWAGLLSPRGSRPWGPTDPRRSRNLG 120  
  
QY 121 KVITDTTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
121 KVITDTTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
  
QY 181 LLSCLTVPASA 191  
DB ||||||||||||  
181 LLSCLTVPASA 191  
  
RESULT 12  
ADE39661  
ID ADE39661 standard; protein; 190 AA.  
XX



```

AC ADE39661;
XX
DT 29-JAN-2004 (first entry)
DE HCV Capsid protein encoded by clone 56.
XX
XX Hepatitis C virus core protein full-length amino acid sequence.
XX
XX hepatitis C virus; HCV; HCV core protein; hepatitis C virus core protein;
KW transfection-associated hepatitis; community-acquired hepatitis;
KW hepatitis; core antigen; anti-core antibody; HCV infection.
XX
XX Hepatitis C virus.
XX
XX EPI310796-A2.
XX
XX 14-MAY-2003.
XX
XX 05-NOV-2002; 2002EP-00257659.
XX
XX 07-NOV-2001; 2001US-0347943P.
XX
XX 10-OCT-2002; 2002US-00268562.
XX
XX (ORTH ) ORTHO CLINICAL DIAGNOSTICS INC.
XX
XX Bahl C;
XX
XX WPI; 2003-543593/52.
XX
XX Antibody/antigen composition useful for detecting hepatitis C virus,
XX comprises mixture of hepatitis C virus protein, modified hepatitis C
XX virus core protein and anti-core antibodies.
XX
XX Example 3; SEQ ID NO 1; 8pp; English.
XX
XX This invention relates to an novel antibody/antigen composition which
XX comprises a mixture of hepatitis C virus (HCV) protein and modified HCV
XX core protein and anti-core antibodies that do not recognise HCV core
XX amino acids 10-43. HCV is a leading cause of transfusion-associated and
XX community-acquired hepatitis. The invention is useful in an immunoassay
XX for HCV or for detecting core antigen and anti-core antibodies in serum
XX collected from HCV infected individuals. The present sequence is that of
XX the full-length hepatitis C virus (HCV) core protein which was used
XX during the creation of the composition of the invention.
XX
XX Sequence 190 AA;
XX
XX Query Match 99.6%; Score 1036; DB 7; Length 190;
XX Best Local Similarity 100.0%; Pred. No. 3.5e-90;
XX Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MSTNPKPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSOPRG 60
DB 1 MSTNPKPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSOPRG 60
QY 61 RROQIPKARRPEGRWTWAGPGYPWPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
DB 61 RROQIPKARRPEGRWTWAGPGYPWPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPLVGAAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPAS 190
DB 181 LLSCLTVPAS 190
XX
XX RESULT 13
XX AAR14557
XX ID AAR14557 standard; protein; 215 AA.
XX
XX AC AAR14557;
XX
XX 24-OCT-2003 (revised)
XX
DT 24-JAN-1992 (first entry)
DE HCV Capsid protein encoded by clone 56.
XX
XX non-A, non-B hepatitis; PT-NANBH; liver disease.
XX
XX Hepatitis C virus; Virus.
XX
XX WO9115516-A.
XX
XX 17-OCT-1991.
XX
XX 06-APR-1990; 90US-00505611.
XX
XX 06-APR-1990; 90US-00505611.
XX
XX 09-OCT-1990; 90US-00594854.
XX
XX (GENE-) GENELABS INC.
XX
XX Reyes G, Kim JP, Moeckli R, Simonsen CC;
XX
XX WPI; 1991-325174/44.
XX
XX N-PSDB; AAQ14279.
XX
XX Hepatitis C virus epitope(s) immuno-reactive with HCV infected sera
XX - useful for detection of HCV infections and as HCV vaccine.
XX
XX Claim 14; Page 110; 150pp; English.
XX
XX Putative HCV capsid protein clones were identified by immunoscreening.
XX Clone 56 was sequenced and found to include a run of adenine residues
XX around nucleotides 25 to 34; such sequences are similar to sequences
XX known to promote translation frame-shifting. The mol. wt. of the protein
XX deduced from the open reading frame is 23.5kd. (Updated on 24-OCT-2003 to
XX standardise OS field)
XX
XX Sequence 215 AA;
XX
XX Query Match 99.6%; Score 1036; DB 2; Length 215;
XX Best Local Similarity 99.5%; Pred. No. 4e-90;
XX Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 MSTNPKPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSOPRG 60
DB 1 MGTNPKPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRRGPRLGVRATRTKTSERSOPRG 60
QY 61 RROQIPKARRPEGRWTWAGPGYPWPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
DB 61 RROQIPKARRPEGRWTWAGPGYPWPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPLVGAAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPAS 191
DB 181 LLSCLTVPAS 191
XX
XX RESULT 14
XX AAR14558
XX ID AAR14558 standard; protein; 215 AA.
XX
XX AC AAR14558;
XX
XX 24-OCT-2003 (revised)
XX
XX 24-JAN-1992 (first entry)
XX
XX Polypeptide encoded by modified HCV Capsid protein clone pGEX-CapA.
XX
XX non-A, non-B hepatitis; PT-NANBH; liver disease.
XX
XX Hepatitis C virus; Virus.
XX

```

```

XX PN WO9115516-A.
XX XX
XX PD 17-OCT-1991.
XX XX
XX PF 06-APR-1990; 90US-00505611.
XX XX
XX PR 06-APR-1990; 90US-00505611.
XX XX
XX PR 09-OCT-1990; 90US-00594854.
XX XX
XX PA (GENE-) GENELABS INC.
XX XX
XX PI Reyes G, Kim JP, Moeckli R, Simonsen CC;
XX XX
XX DR WPI; 1991-325174/44.
XX XX
XX DR N-PSDB; AAQ14280.
XX XX
XX PT Hepatitis C virus epitope(s) immuno-reactive with HCV infected sera
XX PT - useful for detection of HCV infections and as HCV vaccine.
XX PS
XX PS Claim 14; Fig 8B; 150pp; English.
XX XX
XX CC Putative HCV capsid protein clones were identified by immunoscreening.
XX CC Clone 56 was sequenced and found to include a run of adenine residues
XX CC around nucleotides 25 to 34; such sequences are similar to sequences
XX CC known to promote translation frameshifting. To improve expression of the
XX CC HCV capsid protein the putative region of frameshifting was modified
XX CC (i.e. adenine residues at the third position of a codon were changed to
XX CC guanines) to give pGEX-CapA. See AAQ14279 for the unmodified clone 56
XX CC sequence. (Updated on 24-OCT-2003 to standardise OS field)
XX XX
XX SQ Sequence 215 AA;

Query Match 99.6%; Score 1036; DB 2; Length 215;
Best Local Similarity 99.5%; Pred. No. 4e-90;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPDQVFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
DB 1 MGTNPKQKKNKNTNRRPDQVFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
QY 61 RRQIPKARRPEGRWTAQPGYPWPLYGNEGCWAGWLLSPRGRPSWGPTDPRRRRNLG 120
DB 61 RRQIPKARRPEGRWTAQPGYPWPLYGNEGCWAGWLLSPRGRPSWGPTDPRRRRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191

RESULT 15
AAR81945
ID AAR81945 standard; protein; 215 AA.
XX AC
XX AC AAR81945;
XX DT 16-OCT-2003 (revised)
XX DT 23-MAR-2003 (revised)
XX DT 28-JUL-1996 (first entry)
XX DE Hepatitis C virus antigen.
XX XX
XX KW Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;
XX KW antibodies; immunophylaxis; sera; serum.
XX OS Hepatitis C virus; (Clone CapA).
XX XX
XX PN US5443965-A.

```

---

```

PD XX 22-AUG-1995.
PF XX 05-APR-1991; 91US-00681703.
XX XX
PR 06-APR-1990; 90US-00505611.
PR 09-OCT-1990; 90US-00594854.
XX XX
PA (GENE-) GENELABS INC.
XX XX
XX PI Moeckli R, Reyes GR, Kim JP;
XX XX
XX DR WPI; 1995-302120/39.
XX XX
XX DR N-PSDB; AAQ98208.
XX XX
XX PT New nucleic acids encoding hepatitis C virus antigens - used to develop
XX PT prods. for detection of HCV-infected sera and prodn. of vaccines and anti
XX PT -HCV antibodies.
XX PS
XX PS Disclosure; Col 71-72; 71pp; English.
XX XX
XX CC Hepatitis C virus (HCV) antigens can be used for detecting HCV infected
XX CC sera and individuals infected with HCV. They can also be used in an anti-
XX CC HCV vaccine or for the production of anti-HCV antibodies which can be
XX CC used for passive immunophylaxis. The antigens consistently identify
XX CC more HCV positive serum samples with a high degree of specificity. See
XX CC AAQ98202-14 and AAR81939-51. (Updated on 25-MAR-2003 to correct Pf
XX CC field.) (Updated on 25-MAR-2003 to correct PR field.) (Updated on 16-OCT-
XX CC 2003 to standardise OS field)
XX XX
XX SQ Sequence 215 AA;

Query Match 99.6%; Score 1036; DB 2; Length 215;
Best Local Similarity 99.5%; Pred. No. 4e-90;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPDQVFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
DB 1 MGTNPKQKKNKNTNRRPDQVFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
QY 61 RRQIPKARRPEGRWTAQPGYPWPLYGNEGCWAGWLLSPRGRPSWGPTDPRRRRNLG 120
DB 61 RRQIPKARRPEGRWTAQPGYPWPLYGNEGCWAGWLLSPRGRPSWGPTDPRRRRNLG 120
QY 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191

Search completed: August 19, 2005, 12:39:28
Job time : 118.5 secs

```







```

RX MEDLINE=94336721; PubMed=8058787;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the core gene of 14 hepatitis C virus
RL Proc. Natl. Acad. Sci. U.S.A. 91:8239-8243(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93376778; PubMed=8396266;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "At least 12 genotypes of hepatitis C virus predicted by sequence
RL analysis of the putative E1 gene of isolates collected worldwide.";
DR EMBL; U10207; AAA21046.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; ICWX.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
DR CHAIN 1 >191 core protein.
FT NON TER 191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKNTNRPPQDVKFGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
DB 1 MSTNPKPKQKNTNRPPQDVKFGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
QY 61 RRQPIKARPEGRWTAQPGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
DB 61 RRQPIKARPEGRWTAQPGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191

RESULT 5
Q68139 PRELIMINARY; PRT; 191 AA.
ID Q68139
AC Q68139
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94336721; PubMed=8058787;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the core gene of 14 hepatitis C virus
RL Proc. Natl. Acad. Sci. U.S.A. 91:8239-8243(1994).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93376778; PubMed=8396266;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "At least 12 genotypes of hepatitis C virus predicted by sequence
RL analysis of the putative E1 gene of isolates collected worldwide.";
DR EMBL; U10222; AAA21061.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; Q8JYS1; ICWX.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
DR Polyprotein.
DR CHAIN 1 >191 core protein.
FT NON TER 191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKNTNRPPQDVKFGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
DB 1 MSTNPKPKQKNTNRPPQDVKFGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
QY 61 RRQPIKARPEGRWTAQPGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
DB 61 RRQPIKARPEGRWTAQPGYWPPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191

RESULT 6
Q68149 PRELIMINARY; PRT; 191 AA.
ID Q68149
AC Q68149
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=111103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94336721; PubMed=8058787;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the core gene of 14 hepatitis C virus
RL Proc. Natl. Acad. Sci. U.S.A. 91:8239-8243(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92279243; PubMed=1317578;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "Sequence analysis of the 5' noncoding region of hepatitis C virus.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:4942-4946(1992).
RN [3]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=93376778; PubMed=8396266;
RA Bukh J., Purcell R.H., Miller R.H.;
RT "At least 12 genotypes of hepatitis C virus predicted by sequence
RL analysis of the putative E1 gene of isolates collected worldwide.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:8234-8238(1993).
DR EMBL; U10232; AAA21071.1; -.
DR PIR; PQ0804; PQ0804.
DR HSSP; O8JYS1; ICWX.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
KW Polyprotein.
FT CHAIN 1 >191 core protein.
FT NON TER 191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
Db 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSPPSGWPTDPRRSRNILG 120
Db 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSPPSGWPTDPRRSRNILG 120

Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 7
Q6QY91 ID Q6QY91 PRELIMINARY; PRT; 191 AA.
AC Q6QY91;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522326; AAS15488.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
FT NON TER 191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
Db 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSPPSGWPTDPRRSRNILG 120
Db 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSPPSGWPTDPRRSRNILG 120

Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 9
Q6QY99 ID Q6QY99 PRELIMINARY; PRT; 191 AA.
AC Q6QY99;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522326; AAS15488.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
FT NON TER 191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
Db 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSPPSGWPTDPRRSRNILG 120
Db 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSPPSGWPTDPRRSRNILG 120

Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Db 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

```

```
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522318; AAS15480.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Qy 61 RRQIPKARPEGRTPAQPGYPWPLYGNECGWAGWLLSPRGRPSWGPTDPRRSRNLG 120
Db 61 RRQIPKARPEGRTPAQPGYPWPLYGNECGWAGWLLSPRGRPSWGPTDPRRSRNLG 120
Qy 121 KVDTLTTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Db 121 KVDTLTTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 10
Q6QYAL ID Q6QYAL PRELIMINARY; PRT; 191 AA.
AC Q6QYAL;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522316; AAS15478.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Qy 61 RRQIPKARPEGRTPAQPGYPWPLYGNECGWAGWLLSPRGRPSWGPTDPRRSRNLG 120
Db 61 RRQIPKARPEGRTPAQPGYPWPLYGNECGWAGWLLSPRGRPSWGPTDPRRSRNLG 120
Qy 121 KVDTLTTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Db 121 KVDTLTTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 11
Q6QYAL ID Q6QYAL PRELIMINARY; PRT; 191 AA.
AC Q6QYAL;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522315; AAS15477.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Qy 61 RRQIPKARPEGRTPAQPGYPWPLYGNECGWAGWLLSPRGRPSWGPTDPRRSRNLG 120
Db 61 RRQIPKARPEGRTPAQPGYPWPLYGNECGWAGWLLSPRGRPSWGPTDPRRSRNLG 120
Qy 121 KVDTLTTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Db 121 KVDTLTTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 12
Q6QYAL ID Q6QYAL PRELIMINARY; PRT; 191 AA.
AC Q6QYAL;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
```



```
DR EMBL; AY522311; AAS15473.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON_TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNTNRPPQDVKPPGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
Db 1 MSTNPKPQKTKNTNRPPQDVKPPGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRRTWAQPGYPWPLYNGECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120
Db 61 RRQIPKARRPEGRRTWAQPGYPWPLYNGECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120

Qy 121 KVDTLTGCFADLMGYIPAVGAPLGGAARALAHGVRVLEGDVNYATGNLPGCSFSIFLLA 180
Db 121 KVDTLTGCFADLMGYIPAVGAPLGGAARALAHGVRVLEGDVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 13
Q6QYA7 PRELIMINARY; PRT; 191 AA.
AC Q6QYA7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522310; AAS15472.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON_TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNTNRPPQDVKPPGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
Db 1 MSTNPKPQKTKNTNRPPQDVKPPGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRRTWAQPGYPWPLYNGECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120
Db 61 RRQIPKARRPEGRRTWAQPGYPWPLYNGECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120

Qy 121 KVDTLTGCFADLMGYIPAVGAPLGGAARALAHGVRVLEGDVNYATGNLPGCSFSIFLLA 180
Db 121 KVDTLTGCFADLMGYIPAVGAPLGGAARALAHGVRVLEGDVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 14
Q6QYB1 PRELIMINARY; PRT; 191 AA.
AC Q6QYB1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522306; AAS15468.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
DR Pfam; PF01543; HCV_capsid; 1.
DR Pfam; PF01542; HCV_core; 1.
FT NON_TER 191_191
SQ SEQUENCE 191 AA; 20766 MW; 4506596F21B4A893 CRC64;

Query Match 98.7%; Score 1026; DB 2; Length 191;
Best Local Similarity 98.4%; Pred. No. 2e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNTNRPPQDVKPPGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60
Db 1 MSTNPKPQKTKNTNRPPQDVKPPGGQIVGGVYLLPRRGPRLGVRAATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRRTWAQPGYPWPLYNGECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120
Db 61 RRQIPKARRPEGRRTWAQPGYPWPLYNGECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120

Qy 121 KVDTLTGCFADLMGYIPAVGAPLGGAARALAHGVRVLEGDVNYATGNLPGCSFSIFLLA 180
Db 121 KVDTLTGCFADLMGYIPAVGAPLGGAARALAHGVRVLEGDVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 15
Q6QYB2 PRELIMINARY; PRT; 191 AA.
AC Q6QYB2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RA Sherman K.E., Qin H., Shire N., Rouster S.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY522305; AAS15467.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV_capsid.
DR InterPro; IPR002521; HCV_core.
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 12:32:15 ; Search time 24 Seconds  
(without alignments)  
765.725 Million cell updates/sec

Title: US-10-770-117-2

Perfect score: 1040

Sequence: 1 MSTNPKQKGNKNTNRPO.....CSPSIFLLALLSCLTPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: Pir1.\*

2: Pir2.\*

3: Pir3.\*

4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1035	99.5	3011	1 GNWVC3	genome polyprotein
2	1026	98.7	3011	1 GNWVCH	genome polyprotein
3	1015	97.6	3010	1 GNWVCJ	genome polyprotein
4	1013	97.4	3011	1 S40770	genome polyprotein
5	1011	97.2	640	2 J01584	genome polyprotein
6	1010	97.1	441	2 S12707	genome polyprotein
7	1009	97.0	189	2 S32740	polyprotein - hepa
8	1006	96.7	782	2 S18031	genome polyprotein
9	1006	96.7	787	2 PNO677	hypothetical prote
10	1006	96.7	3010	1 GNWVTC	genome polyprotein
11	1002	96.3	513	2 PC1284	genome polyprotein
12	997	95.9	369	2 S21471	genome polyprotein
13	995	95.7	782	2 S19876	genome polyprotein
14	993	95.5	513	2 A44150	structural protein
15	993	95.5	3010	1 A45573	genome polyprotein
16	989	95.1	550	2 JH0711	genome polyprotein
17	988	95.0	520	2 J01925	polyprotein - hepa
18	988	95.0	523	2 J01926	polyprotein - hepa
19	987	94.9	3010	1 S18030	genome polyprotein
20	986	94.8	782	2 S19875	genome polyprotein
21	980	94.2	3010	1 GNWVTW	genome polyprotein
22	970	93.3	782	2 S18032	genome polyprotein
23	969	93.2	874	2 J00883	genome polyprotein
24	967	92.0	876	2 PC2219	polypeptide - hepa
25	964	92.7	3033	1 GNWVJ8	genome polyprotein
26	956	91.9	178	2 P50388	genome polyprotein
27	955	91.8	322	2 JN0265	genome polyprotein
28	951	91.4	874	2 J00881	genome polyprotein
29	951	91.4	3033	1 JQ1303	genome polyprotein

#### ALIGNMENTS

##### RESULT 1

GNWVC3

genome polyprotein - hepatitis C virus (strain HCV-1)

N;Contains: capsid protein C; envelope protein M; hepacivirin (EC 3.4.21.98) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C;Species: hepatitis C virus

C;Date: 30-Sep-1992 #sequence revision 30-Sep-1992 #text\_change 09-Jul-2004

C;Accession: A39166; PQ0403; PQ0404

R;Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, C.; Coit; Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991

A;Title: Genetic organization and diversity of the hepatitis C virus.

A;Reference number: A39166; MUID:91172826; PMID:1848704

A;Accession: A39166

A;Molecule type: mRNA

A;Residues: 1-3011 <CHO>

A;Cross-references: UNIPROT:P26664; GB:M62321; NID:G329873; PIDN:AAA45676.1; PID:G329874

J. Gen. Virol. 73, 1131-1141, 1992

A;Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship to e

A;Reference number: PQ0393; MUID:92268871; PMID:1316939

A;Accession: PQ0403

A;Molecule type: genomic RNA

A;Residues: 1577-1633 <CHA>

A;Cross-references: DDBJ:D10128

A;Experimental source: isolates E-b16

A;Accession: PQ0404

A;Status: preliminary

A;Molecule type: genomic RNA

A;Residues: 1577-1633 <CH2>

A;Experimental source: isolates E-b17

C;Superfamily: hepatitis C virus genome polyprotein

C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura

F;1-115/Product: capsid protein C #status predicted <CPC>

F;116-191/Product: envelope protein M #status predicted <BPM>

F;192-389/Product: major envelope protein E #status predicted <MEE>

F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F;1007-1615/Product: hepacivirin #status predicted <NS3>

F;1230-1237/Region: nucleotide-binding motif A (P-loop)

F;1312-1317/Region: nucleotide-binding motif B

F;1316-1319/Region: DEXH motif

F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>

F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>

F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2077,22

Query Match 99.5% Score 1035; DB 1; Length 3011;

Best Local Similarity 99.5%; Pred. No. 9.6e-77;

Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy

1 MSTNPKQKGNKNTNRPPQVPGGGQIVGGVYLLPRRGPRLGVATRTKTSRSPRG 60  
|||||

```
Db      1  MSTNPKPKKKNTNRRPDQVFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Qy      61  RRQPIPKARRPEGRRTWAQGYPMPLYGNECGWAGWLLSPRGRSPSGWPTDPRRRSRNLG 120
        |||
Db      61  RRQPIPKARRPEGRRTWAQGYPMPLYGNECGWAGWLLSPRGRSPSGWPTDPRRRSRNLG 120
Qy      121  KVDTLTTCGFADLMGYIPAVGAPLGGARALAHGVRVLEGVNYATGNLPGCSFSIFLLA 180
        |||
Db      121  KVDTLTTCGFADLMGYIPLVGAPLGGARALAHGVRVLEGVNYATGNLPGCSFSIFLLA 180
Qy      181  LLSCLTVPASA 191
        |||
Db      181  LLSCLTVPASA 191

RESULT 2
GNMVCH
genome polyprotein - hepatitis C virus (strain H)
N;Contains: capsid protein C; envelope protein M; hepatitis virus (EC 3.4.21.98) (nonstructural
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C;Accession: A36814; A41546
R;Inchauspe, G.; Zebedes, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
submitted to GenBank, July 1992
A;Description: Genomic structure of the human prototype strain H of hepatitis C virus: c
A;Reference number: A36814
A;Accession: A36814
A;Molecule type: Genomic RNA
A;Residues: 1-3011 <INC>
A;Cross-references: UNIPROT:P27958; GB:M67463; NID:G329737; PIDN:AAA45534.1; PID:G329738
R;Inchauspe, G.; Zebedes, S.; Lee, D.H.; Sugitani, M.; Nasoff, M.; Prince, A.M.
Proc. Natl. Acad. Sci. U.S.A. 88, 10292-10296, 1991
A;Title: Genomic structure of the human prototype strain H of hepatitis C virus: compar
A;Reference number: A41546; MUID:92052256; PMID:1658800
A;Contents: annotation
A;Note: neither amino acid nor nucleotide sequence is given
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructura
F;1-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <EPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F;1007-1615/Product: hepatitis virus #status predicted <NS3>
F;1230-1237/Region: nucleotide-binding motif A (P-loop)
F;1312-1317/Region: nucleotide-binding motif B
F;1316-1319/Region: DEXH motif
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4b>
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>
F;196,209,234,305,325,417,423,430,448,476,532,540,556,576,623,645,1213,1255,2041,2240,23

Query Match 98.78; Score 1026; DB 1; Length 3011;
Best Local Similarity 98.4%; Pred. No. 5.2e-76;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy      1  MSTNPKPKKKNTNRRPDQVFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Db      1  MSTNPKPKKKNTNRRPDQVFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60
Qy      61  RRQPIPKARRPEGRRTWAQGYPMPLYGNECGWAGWLLSPRGRSPSGWPTDPRRRSRNLG 120
Db      61  RRQPIPKARRPEGRRTWAQGYPMPLYGNECGWAGWLLSPRGRSPSGWPTDPRRRSRNLG 120
Qy      121  KVDTLTTCGFADLMGYIPAVGAPLGGARALAHGVRVLEGVNYATGNLPGCSFSIFLLA 180
Db      121  KVDTLTTCGFADLMGYIPLVGAPLGGARALAHGVRVLEGVNYATGNLPGCSFSIFLLA 180
Qy      181  LLSCLTVPASA 191
Db      181  LLSCLTVPASA 191
```

RESULT 3  
GNVVCJ

genome polyprotein - hepatitis C virus (strain J)  
N;Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural  
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: hepatitis C virus  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C;Accession: A39253; PS0086  
R;Kato, N.; Hijikata, M.; Ohtsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimoto  
proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990  
A;Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients v  
A;Reference number: A39253; MUID:91088550; PMID:2175903  
A;Accession: A39253  
A;Molecule type: Genomic RNA  
A;Residues: 1-3010 <KAT>  
A;Cross-references: UNIPROT:P26662; GB:D90208; NID:G221610; PIDN:BAAL4233.1; PID:G221611  
R;Kato, N.; Ohkoshi, S.; Shimotohno, K.  
Proc. Jpn. Acad. 65B, 219-223, 1989  
A;Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari  
A;Reference number: PS0085  
A;Accession: PS0085  
A;Molecule type: Genomic RNA  
A;Residues: 2650-2707 <KA2>  
A;Experimental source: Japanese isolate  
A;Comment: The cleavage sites of this polyprotein have not been determined.  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine  
F;2-115/Product: capsid protein C #status predicted <CPC>  
F;116-191/Product: envelope protein M #status predicted <EPM>  
F;192-389/Product: major envelope protein E #status predicted <MEE>  
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F;1007-1615/Product: hepatitis virus #status predicted <NS3>  
F;1230-1237/Region: nucleotide-binding motif A (P-loop)  
F;1312-1317/Region: nucleotide-binding motif B  
F;1316-1319/Region: DEXH motif  
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4a>  
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4b>  
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>  
F;196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,2

Query Match 97.6%; Score 1015; DB 1; Length 3010;  
Best Local Similarity 97.4%; Pred. No. 4.2e-75;  
Matches 186; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MSTNPKPKKKNTNRRPDQVFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60

Db 1 MSTNPKPKKKNTNRRPDQVFPGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRG 60

Qy 61 RRQPIPKARRPEGRRTWAQGYPMPLYGNECGWAGWLLSPRGRSPSGWPTDPRRRSRNLG 120

Db 61 RRQPIPKARRPEGRRTWAQGYPMPLYGNECGWAGWLLSPRGRSPSGWPTDPRRRSRNLG 120

Qy 121 KVDTLTTCGFADLMGYIPAVGAPLGGARALAHGVRVLEGVNYATGNLPGCSFSIFLLA 180

Db 121 KVDTLTTCGFADLMGYIPLVGAPLGGARALAHGVRVLEGVNYATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191

Db 181 LLSCLTVPASA 191

RESULT 4  
S40770

genome polyprotein - hepatitis C virus  
N;Contains: capsid protein C; envelope protein M; hepatitis virus (EC 3.4.21.98) (nonstructu  
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: hepatitis C virus  
C;Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004  
C;Accession: S40770; PC1285  
R;Okamoto, H.

submitted to the EMBL Data Library, March 1992

A;Reference number: S40770  
A;Accession: S40770  
A;Molecule type: genomic RNA  
A;Residues: 1-3011 <OKA>  
A;Cross-references: UNIPROT:Q03463; EMBL:D10749; NID:G221586; PIDN:BAA01582.1; PID:G221586  
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda, Jpn. J. Exp. Med. 60, 167-177, 1990  
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.  
A;Reference number: F1284; MUID:9103116; PMID:21170712  
A;Accession: F1285  
A;Molecule type: Genomic RNA  
A;Residues: 1-513 <OK2>  
A;Cross-references: GB:D00831; NID:G221511; PIDN:BAA00705.1; PID:G221512  
A;Experimental source: isolate HC-J1  
C;Superfamily: hepatitis C virus genome polypeptide  
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polypeptide; serine  
F;2-115/Product: capsid protein C #status predicted <CPC>  
F;116-191/Product: envelope protein M #status predicted <EPM>  
F;192-389/Product: major envelope protein E #status predicted <MEE>  
F;390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F;1007-1615/Product: hepatitis virus #status predicted <NS3>  
F;1230-1237/Region: nucleotide-binding motif A (P-loop)  
F;1312-1317/Region: nucleotide-binding motif B  
F;1316-1319/Region: DEXH motif  
F;1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>  
F;1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>  
F;2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match 97.4%; Score 1013; DB 1; Length 3011;  
Best Local Similarity 97.4%; Pred. No. 6.1e-75;  
Matches 186; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRPPQDVKPPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60

Db 1 MSTIPKQKTKNTNRPPQDVKPPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60

Qy 61 RRQPIKARRPEGRTPAQGYPMWLYGNECGWAGWLLSPRSPSPSGPTDPRRSNGLG 120

Db 61 RRQPIKARRPEGRTPAQGYPMWLYGNECGWAGWLLSPRSPSPSGPTDPRRSNGLG 120

Qy 121 KVDTLTGCFADLMGYIPVAGCAPLGGAARALAHGVRVLEGVNATGNLPCGCSFSIFLLA 180

Db 121 KVDTLTGCFADLMGYIPVAGCAPLGGAARALAHGVRVLEGVNATGNLPCGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191

Db 181 LLSCLTVPASA 191

#### RESULT 5

JQ1584  
genome polypeptide - hepatitis C virus (strain U.K.) (fragment)  
N;Contains: core protein C; envelope protein E1; envelope protein E2; nonstructural protein  
C;Species: hepatitis C virus  
C;Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 09-Jul-2004  
C;Accession: JQ1584  
R;Kumar, U.; Cheng, D.; Thomas, H.; Monjardino, J.  
J. Gen. Virol. 73, 1521-1525, 1992

A;Title: Cloning and sequencing of the structural region and expression of putative core  
A;Reference number: JQ1584; MUID:92300349; PMID:1318944  
A;Accession: JQ1584  
A;Molecule type: Genomic RNA

A;Residues: 1-640 <KUM>  
A;Cross-references: UNIPROT:Q68966; GB:X84079; NID:G643119; PIDN:CAA58888.1; PID:G643120  
C;Superfamily: hepatitis C virus genome polypeptide  
C;Keywords: core protein; envelope protein; glycoprotein; nonstructural protein; polypeptide  
F;1-191/Product: core protein C #status predicted <CPC>  
F;192-389/Product: envelope protein E1 #status predicted <EEL>  
F;390-640/Product: envelope protein E2 and nonstructural protein NS1 #status predicted <NS1>  
F;196, 209, 234, 305, 417, 430, 448, 476, 540, 556, 576, 623/Binding site: carbohydrate (Asn) (cova

Query Match 97.2%; Score 1011; DB 2; Length 640;  
Best Local Similarity 96.9%; Pred. No. 2.1e-75;  
Matches 185; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRPPQDVKPPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60

Db 1 MSTNPKQKQTKNTNRPPQDVKPPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60

Qy 61 RRQPIKARRPEGRTPAQGYPMWLYGNECGWAGWLLSPRSPSPSGPTDPRRSNGLG 120

Db 61 RRQPIKARRPEGRTPAQGYPMWLYGNECGWAGWLLSPRSPSPSGPTDPRRSNGLG 120

Qy 121 KVDTLTGCFADLMGYIPVAGCAPLGGAARALAHGVRVLEGVNATGNLPCGCSFSIFLLA 180

Db 121 KVDTLTGCFADLMGYIPVAGCAPLGGAARALAHGVRVLEGVNATGNLPCGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191

Db 181 LLSCLTVPASA 191

#### RESULT 6

S12707

genome polypeptide - hepatitis C virus (fragment)

N;Contains: core protein; envelope protein

C;Species: hepatitis C virus

C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004

C;Accession: S12707

R;Takeuchi, K.; Kubo, Y.; Boonmar, S.; Watanabe, Y.; Katayama, T.; Choo, Q.-L.; Kuo, G.;

Nucleic Acids Res. 18, 4626, 1990

A;Title: Nucleotide sequence of core and envelope genes of the hepatitis C virus genome

A;Reference number: S12707; MUID:90356432; PMID:2117749

A;Accession: S12707

A;Molecule type: genomic RNA

A;Residues: 1-441 <TAK>

A;Cross-references: UNIPROT:Q81776; EMBL:D00574; NID:G221656; PIDN:BAA00452.1; PID:G221656

C;Superfamily: hepatitis C virus genome polypeptide

C;Keywords: polypeptide

Query Match 97.1%; Score 1010; DB 2; Length 441;  
Best Local Similarity 96.9%; Pred. No. 1.8e-75;  
Matches 185; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRPPQDVKPPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60

Db 1 MSTNPKQKQTKNTNRPPQDVKPPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60

Qy 61 RRQPIKARRPEGRTPAQGYPMWLYGNECGWAGWLLSPRSPSPSGPTDPRRSNGLG 120

Db 61 RRQPIKARRPEGRTPAQGYPMWLYGNECGWAGWLLSPRSPSPSGPTDPRRSNGLG 120

Qy 121 KVDTLTGCFADLMGYIPVAGCAPLGGAARALAHGVRVLEGVNATGNLPCGCSFSIFLLA 180

Db 121 KVDTLTGCFADLMGYIPVAGCAPLGGAARALAHGVRVLEGVNATGNLPCGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191

Db 181 LLSCLTVPASA 191

#### RESULT 7

S32740

polypeptide - hepatitis C virus (isolate Russian) (fragment)

N;Contains: capsid protein C; envelope protein M

C;Species: hepatitis C virus

C;Date: 19-Mar-1997 #sequence\_revision 05-Dec-1998 #text\_change 09-Jul-2004

C;Accession: S32740

R;Vasilev, V.B.; Vazov, S.O.; Kotova, E.Y.; Nosikov, V.V.

submitted to the EMBL Data Library, April 1993

A;Description: Evidence of new HCV variant of European isolate in Russia.

A;Reference number: S32740

A;Accession: S32740

A;Molecule type: genomic RNA

A;Residues: 1-189 <VAS>  
A;Cross-references: UNIPROT:Q68873; EMBL:X71407  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: capsid protein; envelope protein; polyprotein  
F;1-115/Product: capsid protein C #status predicted <CPC>  
F;116-189/Product: envelope protein M #status predicted <EPM>  
  
Query Match 97.0%; Score 1009; DB 2; Length 189;  
Best Local Similarity 97.9%; Pred. No. 9.7e-76;  
Matches 185; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 MSTNPKPKQKXKNTNRRPDQVFPGGQIVGGVYLLPRRGLGVRATKTSERSQPRG 60  
DB 1 MSTNPKPQRTKXNTNRRPDQVFPGGQIVGGVYLLPRRGLGVRATKTSERSQPRG 60  
  
QY 61 RRQPIPKARPEGRRTWAQPGYWPPLYGNECGWAGLLSPRGRSPSWGPTDPRRRSRNLG 120  
DB 61 RRQPIPKARPEGRRTWAQPGYWPPLYGNECGWAGLLSPRGRSPSWGPTDPRRRSRNLG 120  
  
QY 121 KVDTLTTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
DB 121 KVDTLTTCGFADLMGYIPVGLGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
  
QY 181 LLSCLTVPA 189  
DB 181 LLSCLTVPA 189  
  
RESULT 8  
S18031  
genome polyprotein - hepatitis C virus (isolate JK2) (fragment)  
N;Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein  
C;Species: hepatitis C virus  
A;Variety: isolate JK2  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.  
submitted to the EMBL Data Library, September 1991  
A;Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso  
A;Reference number: S18029  
A;Accession: S18031  
A;Molecule type: Genomic RNA  
A;Residues: 1-782 <HON>  
A;Cross-references: UNIPROT:Q68950; EMBL:X61593  
A;Experimental source: isolate JK2  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural  
F;1-191/Product: core protein #status predicted <MAT1>  
F;192-383/Product: envelope protein 1 #status predicted <MAT2>  
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>  
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>  
  
Query Match 96.7%; Score 1006; DB 2; Length 782;  
Best Local Similarity 96.3%; Pred. No. 6.5e-75;  
Matches 184; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 MSTNPKPKQKXKNTNRRPDQVFPGGQIVGGVYLLPRRGLGVRATKTSERSQPRG 60  
DB 1 MSTNPKPQRTKXNTNRRPDQVFPGGQIVGGVYLLPRRGLGVRATKTSERSQPRG 60  
  
QY 61 RRQPIPKARPEGRRTWAQPGYWPPLYGNECGWAGLLSPRGRSPSWGPTDPRRRSRNLG 120  
DB 61 RRQPIPKARPEGRRTWAQPGYWPPLYGNECGWAGLLSPRGRSPSWGPTDPRRRSRNLG 120  
  
QY 121 KVDTLTTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
DB 121 KVDTLTTCGFADLMGYIPVGLGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
  
QY 181 LLSCLTVPA 191  
DB 181 LLSCLTVPA 191  
  
Query Match 96.7%; Score 1006; DB 1; Length 3010;  
Best Local Similarity 96.9%; Pred. No. 2.3e-74;

RESULT 9  
PN0677  
hypothetical protein 787 - hepatitis C virus (fragment)  
C;Species: hepatitis C virus  
C;Date: 14-Jul-1994 #sequence\_revision 14-Jul-1994 #text\_change 09-Jul-2004  
C;Accession: PN0677  
R;Cho, S.H.; Yoon, J.I.; Chang, J.E.; Ahn, B.M.; Lee, C.H.; Lee, Y.I.  
Biochem. Biophys. Res. Commun. 196, 780-788, 1993  
A;Title: Genomic typing of hepatitis C viruses from Korean patients: Implications of gen  
A;Reference number: PN0677; MUID:94059104; PMID:8240354  
A;Accession: PN0677  
A;Molecule type: mRNA  
A;Residues: 1-787 <CHO>  
A;Cross-references: UNIPROT:Q08244; GB:L20498; NID:gl381031; PIDN:AAB02608.1; PID:gl3810  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: glycoprotein; nonstructural protein  
F;196,209,234,250,305,325,421,427,452,536,544,560,580,627,649/Binding site: carbohydrate  
  
Query Match 96.7%; Score 1006; DB 2; Length 787;  
Best Local Similarity 96.3%; Pred. No. 6.5e-75;  
Matches 184; Conservative 3; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 MSTNPKPKQKXKNTNRRPDQVFPGGQIVGGVYLLPRRGLGVRATKTSERSQPRG 60  
DB 1 MSTNPKPQRTKXNTNRRPDQVFPGGQIVGGVYLLPRRGLGVRATKTSERSQPRG 60  
  
QY 61 RRQPIPKARPEGRRTWAQPGYWPPLYGNECGWAGLLSPRGRSPSWGPTDPRRRSRNLG 120  
DB 61 RRQPIPKARPEGRRTWAQPGYWPPLYGNECGWAGLLSPRGRSPSWGPTDPRRRSRNLG 120  
  
QY 121 KVDTLTTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
DB 121 KVDTLTTCGFADLMGYIPVGLGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180  
  
QY 181 LLSCLTVPA 191  
DB 181 LLSCLTVPA 191  
  
RESULT 10  
GNMVT  
genome polyprotein - hepatitis C virus  
N;Contains: capsid protein C; envelope protein M; hepatitis C virus genome isolated from human  
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C;Species: hepatitis C virus  
C;Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004  
C;Accession: A38465  
R;Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.;  
J. Virol. 65, 1105-1113, 1991  
A;Title: Structure and organization of the hepatitis C virus genome  
A;Reference number: A38465; MUID:91140698; PMID:1847440  
A;Accession: A38465  
A;Molecule type: Genomic RNA  
A;Residues: 1-3010 <TAK>  
A;Cross-references: UNIPROT:P26663; EMBL:M58335; NID:g329770; PIDN:AAA72945.1; PID:g3297;  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural  
F;2-115/Product: capsid protein C #status predicted <CPC>  
F;116-191/Product: envelope protein M #status predicted <EPM>  
F;192-389/Product: major envelope protein NS1 #status predicted <NS1>  
F;390-729/Product: nonstructural protein NS2 #status predicted <NS2>  
F;730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F;1007-1615/Product: hepatitis C virus genome isolated from human  
F;1230-1237/Region: nucleotide-binding motif A (P-loop)  
F;1312-1317/Region: nucleotide-binding motif B  
F;1316-1319/Region: DEH motif  
F;1863-2013/Product: nonstructural protein NS4a #status predicted <N4a>  
F;2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>  
F;196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,224  
  
Query Match 96.7%; Score 1006; DB 1; Length 3010;  
Best Local Similarity 96.9%; Pred. No. 2.3e-74;

Matches 185; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 MSTNPKQKKNTNRRPQDVKPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
Db |||||:|||||  
Qy 1 MSTNPKFQRTKNTNRRPQDVKPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
Db |||||:|||||  
Qy 61 RRQPIPKARPEGRGRTWAQPGYPWPLYGNEGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Db 61 RRQPIPKARPEGRGRTWAQPGYPWPLYGNEGLWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPLVGAAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTTPASA 191

RESULT 11  
PC1284  
genome polyprotein - hepatitis C virus (isolate HC-J4) (fragment)  
N;Contains: hepatitis C virus  
C;Species: hepatitis C virus  
C;Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C;Accession: PC1284  
R;Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Teuda, Jpn. J. Exp. Med. 60, 167-177, 1990  
A;Title: The 5'-terminal sequence of the hepatitis C virus genome.  
A;Reference number: PC1284; MUID:91013116; PMID:2170712  
A;Accession: PC1284  
A;Molecule type: genomic RNA  
A;Residues: 1-513 <OKA>  
A;Cross-references: UNIPROT:Q81221; GB:D00832; NID:g221513; PIDN:BAA00706.1; PID:g221514  
C;Superfamily: hepatitis C virus genome polyprotein

Query Match 96.3%; Score 1002; DB 2; Length 513;  
Best Local Similarity 96.3%; Pred. No. 9.3e-75;  
Matches 184; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 MSTNPKQKKNTNRRPQDVKPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
Db |||||:|||||  
Qy 61 RRQPIPKARPEGRGRTWAQPGYPWPLYGNEGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Db 61 WRQPIPKARPEGRGRTWAQPGYPWPLYGNEGLWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPLVGAAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTTPASA 191

RESULT 12  
S21471  
genome polyprotein - hepatitis C virus (fragment)  
N;Contains: capsid protein; envelope protein  
C;Species: hepatitis C virus  
C;Date: 20-Feb-1995 #sequence\_revision 19-Apr-1996 #text\_change 09-Jul-2004  
C;Accession: S21471  
R;Mogam, W.K.  
A;Reference number: S21471  
A;Accession: S21471  
A;Molecule type: genomic RNA  
A;Residues: 1-369 <MOG>  
A;Cross-references: UNIPROT:Q68869; EMBL:X65924; NID:g59466; PIDN:CAA46717.1; PID:g59467  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: capsid protein; envelope protein; glycoprotein; polyprotein

Query Match 95.9%; Score 997; DB 2; Length 369;  
Best Local Similarity 95.8%; Pred. No. 1.8e-74;  
Matches 183; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 MSTNPKQKKNTNRRPQDVKPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
Db |||||:|||||  
Qy 1 MSTNPKFQRTKNTNRRPQDVKPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
Db |||||:|||||  
Qy 61 RRQPIPKARPEGRGRTWAQPGYPWPLYGNEGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Db 61 RRQPIPKARPEGRGRTWAQPGYPWPLYGNEGLWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPLVGAAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTTPASA 191

RESULT 13  
S19876  
genome polyprotein - hepatitis C virus (isolate JK5) (fragment)  
N;Contains: core protein; envelope protein 1; nonstructural protein 2; NS1/E2 protein  
C;Species: hepatitis C virus  
A;Variety: isolate JK5  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C;Accession: S19876  
R;Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S. submitted to the EMBL Data Library, September 1991  
A;Description: Sequence analysis of putative structural regions of Hepatitis C Virus iso A;Reference number: S18029  
A;Accession: S19876  
A;Molecule type: genomic RNA  
A;Residues: 1-782 <HON>  
A;Cross-references: UNIPROT:Q68953; EMBL:X61595; NID:g59486; PIDN:CAA43792.1; PID:g59487  
A;Experimental source: isolate JK5  
C;Superfamily: hepatitis C virus genome polyprotein  
C;Keywords: capsid protein; core protein; envelope protein; glycoprotein; nonstructural F;1-191/Product: core protein #status predicted <MAT1>  
F;192-383/Product: envelope protein 1 #status predicted <MAT2>  
F;384-733/Product: NS1/E2 protein #status predicted <MAT3>  
F;734-782/Product: nonstructural protein 2 (fragment) #status predicted <MAT4>

Query Match 95.7%; Score 995; DB 2; Length 782;  
Best Local Similarity 95.3%; Pred. No. 5.2e-74;  
Matches 182; Conservative 3; Mismatches 6; Indels 0; Gaps 0;  
Qy 1 MSTNPKQKKNTNRRPQDVKPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
Db 1 MSTNPKFQRTKNTNRRPQDVKPGGQIVGGVYLLPRGPRGLGVRAATKTSERSQPRG 60  
Qy 61 RRQPIPKARPEGRGRTWAQPGYPWPLYGNEGCGWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Db 61 RRQPIPKARPEGRGRTWAQPGYPWPLYGNEGLWAGWLLSPRGRSPSWGPTDPRRRSRNLG 120  
Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPLVGAAPLGGAAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTTPASA 191

RESULT 14  
A44150  
structural protein - hepatitis C virus  
C;Species: hepatitis C virus  
C;Date: 27-Jun-1994 #sequence\_revision 27-Jun-1994 #text\_change 09-Jul-2004  
C;Accession: A44150  
R;Ching, W.M.; Wychowski, C.; Beach, M.J.; Wang, H.; Davies, C.L.; Carl, M.; Bradley, D.







GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 12:43:21 ; Search time 107 Seconds  
(without alignments)

698.999 Million cell updates/sec

Title: US-10-770-117-2

Perfect score: 1040

Sequence: 1 MSTNPKQKKNKENTNRPPQ.....CSPSIFLLALLSCLTVPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US11B\_PUBCOMB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1040	100.0	191	18	US-10-770-117-2
2	1035	99.5	191	18	US-10-770-117-4
3	1035	99.5	3011	9	US-09-916-359-2
4	1035	99.5	3011	16	US-10-445-724-2
5	1031	99.1	190	14	US-10-268-562-1
6	1030	99.0	190	15	US-10-450-649-7
7	1026	98.7	249	15	US-10-365-620-54
8	1026	98.7	249	17	US-10-912-969-56
9	1026	98.7	450	15	US-10-651-165-179
10	1026	98.7	450	15	US-10-651-165-180
11	1026	98.7	473	15	US-10-365-620-56

12	1026	98.7	473	17	US-10-912-969-58	Sequence 58, Appl
13	1026	98.7	473	17	US-10-913-171-39	Sequence 39, Appl
14	1026	98.7	3011	9	US-09-742-659-4	Sequence 4, Appl
15	1026	98.7	3011	9	US-09-952-572-9	Sequence 9, Appl
16	1026	98.7	3011	9	US-09-929-955-1	Sequence 1, Appl
17	1026	98.7	3011	9	US-09-747-419-20	Sequence 20, Appl
18	1026	98.7	3011	10	US-09-891-894-3	Sequence 3, Appl
19	1026	98.7	3011	13	US-10-104-966-1	Sequence 1, Appl
20	1026	98.7	3011	14	US-10-259-275-20	Sequence 20, Appl
21	1026	98.7	3011	14	US-10-184-150-3	Sequence 3, Appl
22	1026	98.7	3011	15	US-10-328-997-3	Sequence 3, Appl
23	1026	98.7	3011	15	US-10-189-359-14	Sequence 14, Appl
24	1026	98.7	3011	15	US-10-296-734-406	Sequence 406, Appl
25	1026	98.7	3011	15	US-10-719-619-1	Sequence 1, Appl
26	1026	98.7	3011	16	US-10-817-591-1	Sequence 1, Appl
27	1026	98.7	3011	20	US-11-006-313-20	Sequence 20, Appl
28	1026	98.7	3012	9	US-09-238-076-2	Sequence 2, Appl
29	1026	98.7	3012	10	US-09-995-937-2	Sequence 2, Appl
30	1026	98.7	3012	10	US-09-917-563-2	Sequence 2, Appl
31	1022	98.3	3011	9	US-09-238-076-20	Sequence 20, Appl
32	1022	98.3	3011	10	US-09-995-937-20	Sequence 20, Appl
33	1022	98.3	3011	10	US-09-917-563-20	Sequence 20, Appl
34	1015	97.6	450	15	US-10-651-165-187	Sequence 187, App
35	1013	97.4	450	15	US-10-651-165-181	Sequence 181, App
36	1013	97.4	2894	9	US-09-941-611-23	Sequence 23, Appl
37	1013	97.4	2894	14	US-10-044-995-23	Sequence 23, Appl
38	1013	97.4	2894	17	US-10-822-871-23	Sequence 23, Appl
39	1011	97.2	450	15	US-10-651-165-190	Sequence 190, App
40	1011	97.2	3011	14	US-10-232-643-6	Sequence 6, Appl
41	1007	96.8	809	9	US-09-973-025-50	Sequence 50, Appl
42	1007	96.8	809	10	US-09-899-303-50	Sequence 50, Appl
43	1007	96.8	809	10	US-09-995-808-50	Sequence 50, Appl
44	1007	96.8	809	10	US-09-995-860-50	Sequence 50, Appl
45	1007	96.8	809	10	US-09-995-791-50	Sequence 50, Appl

#### ALIGNMENTS

RESULT 1  
US-10-770-117-2  
; Sequence 2, Application US/10770117  
; Publication No. US20050129705A1  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States, as represented by the  
; APPLICANT: Secretary, Dept. of  
; APPLICANT: Health and Human Services  
; APPLICANT: Berzofsky, Jay A.  
; APPLICANT: Sarobe, Pablo  
; APPLICANT: Pendleton, C. David  
; APPLICANT: Reinstone, Stephen M.  
; APPLICANT: Arichi, Tatsumi  
; APPLICANT: Major, Marian E.  
; TITLE OF INVENTION: MODIFIED HCV PEPTIDE VACCINES  
; FILE REFERENCE: 14014.0347/P  
; CURRENT APPLICATION NUMBER: US/10/770,117  
; CURRENT FILING DATE: 2004-02-02  
; PRIOR APPLICATION NUMBER: US/09/763,260  
; PRIOR FILING DATE: 2001-10-19  
; PRIOR APPLICATION NUMBER: 60/097,446  
; PRIOR FILING DATE: 1998-08-21  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/Note =  
; OTHER INFORMATION: synthetic construct  
US-10-770-117-2

Query Match 100.0%; Score 1040; DB 18; Length 191;

```
Best Local Similarity 100.0%; Pred. No. 5.2e-80;
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
DB 1 MSTNPKQKKNKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

QY 61 RRQPIPKARPEGRTWAQPGYPWPLYGNECGWAGWLLSPRGRSPSGWPTDPRRSRNLG 120
DB 61 RRQPIPKARPEGRTWAQPGYPWPLYGNECGWAGWLLSPRGRSPSGWPTDPRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191

RESULT 2
US-10-770-117-4
; Sequence 4, Application US/10770117
; Publication No. US20050129705A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States, as represented by the
; APPLICANT: Secretary, Dept. of
; APPLICANT: Health and Human Services
; APPLICANT: Berzofsky, Jay A.
; APPLICANT: Sarobe, Pablo
; APPLICANT: Pendleton, C. David
; APPLICANT: Feinstein, Stephen M.
; APPLICANT: Arichi, Tatsuemi
; APPLICANT: Major, Marian E.
; TITLE OF INVENTION: MODIFIED HCV PEPTIDE VACCINES
; FILE REFERENCE: 14014.0347/P
; CURRENT APPLICATION NUMBER: US/10/770,117
; CURRENT FILING DATE: 2004-02-02
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US/09/763,260
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 60/097,446
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 191
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:/Note =
US-10-770-117-4

Query Match 99.5%; Score 1035; DB 18; Length 191;
Best Local Similarity 99.5%; Pred. No. 1.4e-79;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
DB 1 MSTNPKQKKNKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

QY 61 RRQPIPKARPEGRTWAQPGYPWPLYGNECGWAGWLLSPRGRSPSGWPTDPRRSRNLG 120
DB 61 RRQPIPKARPEGRTWAQPGYPWPLYGNECGWAGWLLSPRGRSPSGWPTDPRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191
```

```
RESULT 3
US-09-916-359-2
; Sequence 2, Application US/09916359
; Patent No. US20020034734A1
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; TITLE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: PMCF97-03A
; CURRENT APPLICATION NUMBER: US/09/916,359
; CURRENT FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 09/388,874
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: 97/02,887
; PRIOR FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Virus
US-09-916-359-2

Query Match 99.5%; Score 1035; DB 9; Length 3011;
Best Local Similarity 99.5%; Pred. No. 2.5e-78;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
DB 1 MSTNPKQKKNKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

QY 61 RRQPIPKARPEGRTWAQPGYPWPLYGNECGWAGWLLSPRGRSPSGWPTDPRRSRNLG 120
DB 61 RRQPIPKARPEGRTWAQPGYPWPLYGNECGWAGWLLSPRGRSPSGWPTDPRRSRNLG 120

QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180

QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191

RESULT 4
US-10-445-724-2
; Sequence 2, Application US/10445724
; Publication No. US20040101829A1
; GENERAL INFORMATION:
; APPLICANT: STAPLETON, JACK T.
; APPLICANT: WUENSCHMANN, SABINA
; TITLE OF INVENTION: A PROTEIN THAT INTERACTS WITH LIPIDS AND
; TITLE OF INVENTION: METHODS FOR TREATING HYPERLIPIDEMIA
; FILE REFERENCE: IOWA:045US
; CURRENT APPLICATION NUMBER: US/10/445,724
; CURRENT FILING DATE: 2003-05-27
; PRIOR APPLICATION NUMBER: 60/392,158
; PRIOR FILING DATE: 2003-05-27
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-445-724-2

Query Match 99.5%; Score 1035; DB 16; Length 3011;
Best Local Similarity 99.5%; Pred. No. 2.5e-78;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
DB 1 MSTNPKQKKNKNTNRRPQDVKFPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60
```

```
Db      1 MSTNPKQKKNKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
Qy      61 RROPIPKARPEGRWTWAPGYWPLYNCGCGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
Db      61 RROPIPKARPEGRWTWAPGYWPLYNCGCGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
Qy      121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Db      121 KVIDTLTCGFADLMGYIPVAGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
Qy      181 LLSCLTVPASA 191
Db      181 LLSCLTVPASA 191

RESULT 5
US-10-268-562-1
; Sequence 1, Application US/10268562
; Publication No. US20030108563A1
; GENERAL INFORMATION:
; APPLICANT: Otho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Reagents for the simultaneous detection of HCV core antigens and
; FILE REFERENCE: antibodies
; CURRENT APPLICATION NUMBER: US/10/268,562
; CURRENT FILING DATE: 2002-10-10
; PRIOR FILING DATE: 2002-10-10
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-268-562-1

Query Match      99.1%; Score 1031; DB 14; Length 190;
Best Local Similarity 99.5%; Pred. No. 3e-79; Indels 0; Gaps 0;
Matches 189; Conservative 0; Mismatches 1;

1 MSTNPKQKKNKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
1 MSTNPKQKKNKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
61 RROPIPKARPEGRWTWAPGYWPLYNCGCGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
61 RROPIPKARPEGRWTWAPGYWPLYNCGCGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
121 KVIDTLTCGFADLMGYIPVAGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
181 LLSCLTVPASA 190
181 LLSCLTVPAS 190

US-10-450-649-7
; Sequence 7, Application US/10450649
; Publication No. US20040052818A1
; GENERAL INFORMATION:
; APPLICANT: Heinz, Franz X.
; TITLE OF INVENTION: ATTENUATED LIVE VACCINE
; FILE REFERENCE: U 014666-0
; CURRENT APPLICATION NUMBER: US/10/450,649
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: PCT/AT02/00046
; PRIOR FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: A 272/2001 AT
; PRIOR FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 9

Query Match      99.1%; Score 1031; DB 14; Length 190;
Best Local Similarity 99.5%; Pred. No. 3e-79; Indels 0; Gaps 0;
Matches 189; Conservative 0; Mismatches 1;

1 MSTNPKQKKNKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
1 MSTNPKQKKNKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
61 RROPIPKARPEGRWTWAPGYWPLYNCGCGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
61 RROPIPKARPEGRWTWAPGYWPLYNCGCGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
121 KVIDTLTCGFADLMGYIPVAGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
181 LLSCLTVPAS 190
181 LLSCLTVPAS 190
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Hepatitis C Virus 1
US-10-450-649-7

Query Match      99.0%; Score 1030; DB 15; Length 190;
Best Local Similarity 99.5%; Pred. No. 3.6e-79; Indels 0; Gaps 0;
Matches 189; Conservative 0; Mismatches 1;

2 STNPKQKKNKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 61
1 STNPKQKKNKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
62 RROPIPKARPEGRWTWAPGYWPLYNCGCGWAGWLLSPRGSRPSWGPTDPRRSRNLG 121
61 RROPIPKARPEGRWTWAPGYWPLYNCGCGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
122 VIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 181
121 VIDTLTCGFADLMGYIPVAGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
182 LLSCLTVPASA 191
181 LLSCLTVPASA 190

RESULT 7
US-10-365-620-54
; Sequence 54, Application US/10365620
; Publication No. US20040001853A1
; GENERAL INFORMATION:
; APPLICANT: George, Rajan
; APPLICANT: Tyrrell, Lorne
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 656.0016
; CURRENT APPLICATION NUMBER: US/10/365,620
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US60/423,578
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/390,564
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 54
; LENGTH: 249
; TYPE: PRT
; ORGANISM: ORF of HCV Core Protein
US-10-365-620-54

Query Match      98.7%; Score 1026; DB 15; Length 249;
Best Local Similarity 98.4%; Pred. No. 1e-78; Indels 2; Gaps 0;
Matches 188; Conservative 1; Mismatches 2;

1 MSTNPKQKKNKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 60
31 MSTNPKQKKNTNRRPQDVKPPGGQIVGGVYLLPRRGPRLGVRATRTKTSERSQPRG 90
61 RROPIPKARPEGRWTWAPGYWPLYNCGCGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120
91 RROPIPKARPEGRWTWAPGYWPLYNCGCGWAGWLLSPRGSRPSWGPTDPRRSRNLG 150
121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180
151 KVIDTLTCGFADLMGYIPVAGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 210
181 LLSCLTVPASA 191
211 LLSCLTVPASA 221
```

```
RESULT 8
US-10-912-969-56
; Sequence 56, Application US/10912969
; Publication No. US20050013828A1
; GENERAL INFORMATION:
; APPLICANT: Virexx Research, Inc.
; APPLICANT: George, Rajan
; APPLICANT: Tyrrell, Lorne
; APPLICANT: Noujaim, Antoine
; APPLICANT: Wang, Dakun
; APPLICANT: Ma, Allan
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 17506-007001
; CURRENT APPLICATION NUMBER: US/10/912,969
; CURRENT FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: US 60/390,564
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/423,578
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 10/365,620
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: PCT/IB04/00373
; PRIOR FILING DATE: 2004-02-14
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 249
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-912-969-56

Query Match          98.7%; Score 1026; DB 17; Length 249;
Best Local Similarity 98.4%; Pred. No. 1e-78; 2; Indels 0; Gaps 0;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKKKNTNRRPQDVKFPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
DB 31 MSTNPKPQKTKNTNRRPQDVKFPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 90
QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECCGAGWLLSPRGRPSWGPTDPRRRSRNLG 120
DB 91 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECCGAGWLLSPRGRPSWGPTDPRRRSRNLG 150
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
DB 151 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRLVDGVNYATGNLPGCSFSIFLLA 210
QY 181 LLSCLTVPASA 191
DB 211 LLSCLTVPASA 221

RESULT 9
US-10-651-165-179
; Sequence 179, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 180
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-179

Query Match          98.7%; Score 1026; DB 15; Length 450;
Best Local Similarity 98.4%; Pred. No. 2e-78; 2; Indels 0; Gaps 0;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKKKNTNRRPQDVKFPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
DB 1 MSTNPKPQKTKNTNRRPQDVKFPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECCGAGWLLSPRGRPSWGPTDPRRRSRNLG 120
DB 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECCGAGWLLSPRGRPSWGPTDPRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191

RESULT 10
US-10-651-165-180
; Sequence 180, Application US/10651165
; Publication No. US20040047877A1
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 180
; LENGTH: 450
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-180

Query Match          98.7%; Score 1026; DB 15; Length 450;
Best Local Similarity 98.4%; Pred. No. 2e-78; 2; Indels 0; Gaps 0;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTNPKPKQKKKNTNRRPQDVKFPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
DB 1 MSTNPKPQKTKNTNRRPQDVKFPGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60
QY 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECCGAGWLLSPRGRPSWGPTDPRRRSRNLG 120
DB 61 RRQPIPKARRPEGRRTWAQPGYPWPLYGNECCGAGWLLSPRGRPSWGPTDPRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPLVGAPLGGAAALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191
```

```
RESULT 11
US-10-365-620-56
; Sequence 56, Application US/10365620
; Publication No. US20040001853A1
; GENERAL INFORMATION:
; APPLICANT: George, Rajan
; APPLICANT: Tyrell, Lorne
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 656.0016
; CURRENT APPLICATION NUMBER: US/10/365,620
; CURRENT FILING DATE: 2003-02-13
; PRIOR FILING DATE: 2003-11-05
; PRIOR APPLICATION NUMBER: 60/390,564
; PRIOR FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 56
; LENGTH: 473
; TYPE: PRT
; ORGANISM: ORF of HCV Core-TED protein
US-10-365-620-56

Query Match          98.7%; Score 1026; DB 15; Length 473;
Best Local Similarity 98.4%; Pred. No. 2.1e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKPKQKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db 31 MSTNPKPKQKTKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 90
Qy 61 RRQPIPKARPEGRRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 120
Db 91 RRQPIPKARPEGRRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 150
Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db 151 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 210
Qy 181 LLSCLTVPASA 191
Db 211 LLSCLTVPASA 221

RESULT 12
US-10-912-969-58
; Sequence 58, Application US/10912969
; Publication No. US20050013828A1
; GENERAL INFORMATION:
; APPLICANT: Virexx Research, Inc.
; APPLICANT: George, Rajan
; APPLICANT: Tyrell, Lorne
; APPLICANT: Noujaim, Antoine
; APPLICANT: Wang, Dakun
; APPLICANT: Ma, Allan
; TITLE OF INVENTION: Chimeric Antigens for Eliciting An Immune Response
; FILE REFERENCE: 17506-007001
; CURRENT APPLICATION NUMBER: US/10/912,969
; CURRENT FILING DATE: 2004-08-05
; PRIOR FILING DATE: 2002-06-20
; PRIOR APPLICATION NUMBER: US 60/390,564
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 10/365,620
; PRIOR FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: PCT/IB04/00373
; PRIOR FILING DATE: 2004-02-14
; NUMBER OF SEQ ID NOS: 79
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 58
```

```
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-912-969-58

Query Match          98.7%; Score 1026; DB 17; Length 473;
Best Local Similarity 98.4%; Pred. No. 2.1e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKPKQKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db 31 MSTNPKPKQKTKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 90
Qy 61 RRQPIPKARPEGRRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 120
Db 91 RRQPIPKARPEGRRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 150
Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db 151 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 210
Qy 181 LLSCLTVPASA 191
Db 211 LLSCLTVPASA 221

RESULT 13
US-10-913-171-39
; Sequence 39, Application US/10913171
; Publication No. US20050031628A1
; GENERAL INFORMATION:
; APPLICANT: Virexx Research, Inc.
; APPLICANT: George, Rajan
; APPLICANT: Tyrell, Lorne
; APPLICANT: Noujaim, Antoine
; APPLICANT: Wang, Dakun
; APPLICANT: Ma, Allan
; TITLE OF INVENTION: CHIMERIC ANTIGENS FOR BREAKING HOST TOLERANCE TO FOREIGN ANTIGENS
; FILE REFERENCE: 17506-006001
; CURRENT APPLICATION NUMBER: US/10/913,171
; CURRENT FILING DATE: 2004-08-05
; PRIOR APPLICATION NUMBER: US 60/493,449
; PRIOR FILING DATE: 2004-08-08
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-913-171-39

Query Match          98.7%; Score 1026; DB 17; Length 473;
Best Local Similarity 98.4%; Pred. No. 2.1e-78;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKPKQKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db 31 MSTNPKPKQKTKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 90
Qy 61 RRQPIPKARPEGRRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 120
Db 91 RRQPIPKARPEGRRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSPWGPTDPRRRSRNLG 150
Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
Db 151 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 210
Qy 181 LLSCLTVPASA 191
Db 211 LLSCLTVPASA 221
```

```
Db      211  LLSCLTVPASA 221

RESULT 14
US-09-742-659-4
; Sequence 4, Application US/09742659
; Patent No. US20010034019A1
; GENERAL INFORMATION:
; APPLICANT: Hong, Zhi
; APPLICANT: Butkiewicz, Nancy J.
; APPLICANT: Zhong, Weidong
; APPLICANT: Ingravallo, Paul
; APPLICANT: Wright-Minogue, Jacquelyn
; APPLICANT: Lau, Johnson Y.
; APPLICANT: Lemon, Stanley M.
; TITLE OF INVENTION: Chimeric HCV/GBV-B viruses
; FILE REFERENCE: ID01116
; CURRENT APPLICATION NUMBER: US/09/742,659
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: US 60/171,469
; PRIOR FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3011
; TYPE: PRN
; ORGANISM: Hepatitis C virus
US-09-742-659-4

Query Match      98.7%; Score 1026; DB 9; Length 3011;
Best Local Similarity 98.4%; Pred. No. 1.4e-77;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1  MSTNPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1  MSTNPKPQKTKKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      61  RRQPIPKARRPEGRTPAQGPWPPLYGNECGWAGWLLSPRGSRRPSWGPTDPRRRSRNLG 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61  RRQPIPKARRPEGRTPAQGPWPPLYGNECGWAGWLLSPRGSRRPSWGPTDPRRRSRNLG 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      121  KVIDLTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121  KVIDLTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      181  LLSCLTVPASA 191
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181  LLSCLTVPASA 191
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: August 19, 2005, 13:01:50
Job time : 108 secs

Query Match      98.7%; Score 1026; DB 9; Length 3011;
Best Local Similarity 98.4%; Pred. No. 1.4e-77;
Matches 188; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1  MSTNPKQKKNKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      1  MSTNPKPQKTKKNTNRRPDQVKFPGGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      61  RRQPIPKARRPEGRTPAQGPWPPLYGNECGWAGWLLSPRGSRRPSWGPTDPRRRSRNLG 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      61  RRQPIPKARRPEGRTPAQGPWPPLYGNECGWAGWLLSPRGSRRPSWGPTDPRRRSRNLG 120
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      121  KVIDLTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      121  KVIDLTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      181  LLSCLTVPASA 191
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      181  LLSCLTVPASA 191
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
US-09-952-572-9
; Sequence 9, Application US/09952572
; Patent No. US20020119495A1
; GENERAL INFORMATION:
; APPLICANT: HAWAII BIOTECHNOLOGY GROUP, Inc.
; APPLICANT: NAKANO, Eileen
; APPLICANT: CLEMENTS, David
; APPLICANT: HUMPHREYS, Tom
; TITLE OF INVENTION: IMMUNOGENIC COMPOSITION OF HEPATITIS C
; FILE REFERENCE: HAWBIO1100
; CURRENT APPLICATION NUMBER: US/09/952,572
; CURRENT FILING DATE: 2001-09-13
; PRIOR APPLICATION NUMBER: US 60/230,927
; PRIOR FILING DATE: 2000-09-13
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 3011
; TYPE: PRN
; ORGANISM: Hepatitis C Virus
US-09-952-572-9
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 12:34:55 ; Search time 20 Seconds  
(without alignments)  
712.899 Million cell updates/sec

Title: US-10-770-117-2

Perfect score: 1040

Sequence: 1 MSTNPKQKKNKNTNRPPQ.....CSFSIFLLALLSCLTVPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents AA:\*
- 1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep:\*
  - 2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep:\*
  - 3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep:\*
  - 4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep:\*
  - 5: /cgn2\_6/ptodata/1/iaa/PTCUS COMB.pep:\*
  - 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1040	100.0	191	4	US-09-763-260-2
2	1035	99.5	191	4	US-09-763-260-4
3	1035	99.5	3011	1	US-08-440-103-36
4	1035	99.5	3011	1	US-08-440-542-36
5	1035	99.5	3011	1	US-07-910-760-10
6	1035	99.5	3011	1	US-08-440-519-10
7	1035	99.5	3011	1	US-08-231-368-36
8	1035	99.5	3011	1	US-08-440-210-36
9	1035	99.5	3011	3	US-09-388-874-2
10	1035	99.5	3011	3	US-09-046-604-36
11	1035	99.5	3011	3	US-08-440-549-10
12	1035	99.5	3011	3	US-08-850-328-1
13	1035	99.5	3011	4	US-09-916-359-2
14	1031	99.1	215	1	US-07-681-703B-12
15	1031	99.1	215	1	US-07-681-703B-14
16	1031	99.1	215	5	PCT-US91-02370-12
17	1031	99.1	215	5	PCT-US91-02370-14
18	1031	99.1	217	2	US-08-407-410B-12
19	1031	99.1	217	2	US-08-407-410B-14
20	1031	99.1	217	2	US-08-485-500-12
21	1031	99.1	217	2	US-08-485-500-14
22	1029	98.9	2995	3	US-08-444-818-138
23	1026	98.7	191	2	US-08-290-665A-156
24	1026	98.7	191	2	US-08-290-665A-157
25	1026	98.7	191	2	US-08-290-665A-158
26	1026	98.7	191	2	US-08-290-665A-159
27	1026	98.7	191	2	US-08-290-665A-160

28	1026	98.7	191	3	US-08-380-160-3	Sequence 3, Appli
29	1026	98.7	191	5	PCT-US95-10398-156	Sequence 156, App
30	1026	98.7	191	5	PCT-US95-10398-157	Sequence 157, App
31	1026	98.7	191	5	PCT-US95-10398-158	Sequence 158, App
32	1026	98.7	191	5	PCT-US95-10398-159	Sequence 159, App
33	1026	98.7	191	5	PCT-US95-10398-160	Sequence 160, App
34	1026	98.7	450	4	US-08-635-886C-179	Sequence 179, App
35	1026	98.7	450	4	US-08-635-886C-180	Sequence 180, App
36	1026	98.7	450	4	US-08-635-886C-179	Sequence 179, App
37	1026	98.7	450	4	US-08-635-886C-180	Sequence 180, App
38	1026	98.7	967	1	US-08-188-281B-13	Sequence 13, Appl
39	1026	98.7	967	5	PCT-US94-07280-13	Sequence 13, Appl
40	1026	98.7	967	5	PCT-US95-01087-13	Sequence 13, Appl
41	1026	98.7	1648	1	US-08-188-281B-12	Sequence 12, Appl
42	1026	98.7	1648	5	PCT-US94-07280-12	Sequence 12, Appl
43	1026	98.7	1648	5	PCT-US95-01087-12	Sequence 12, Appl
44	1026	98.7	3011	1	US-08-188-281B-1	Sequence 1, Appli
45	1026	98.7	3011	1	US-08-453-552-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-763-260-2

; Sequence 2, Application US/09763260

; Patent No. 6685944

; GENERAL INFORMATION:

; APPLICANT: The Government of the United States, as represented by the

; APPLICANT: Secretary, Dept. of

; APPLICANT: Health and Human Services

; APPLICANT: Berzofsky, Jay A.

; APPLICANT: Sarobe, Pablo

; APPLICANT: Pendleton, C. David

; APPLICANT: Feinstein, Stephen M.

; APPLICANT: Arichi, Tatsumi

; APPLICANT: Major, Marian E.

; TITLE OF INVENTION: MODIFIED HCV PEPTIDE VACCINES

; FILE REFERENCE: 14014.0347/P

; CURRENT APPLICATION NUMBER: US/09/763,260

; CURRENT FILING DATE: 2001-10-19

; PRIOR APPLICATION NUMBER: 60/097,446

; PRIOR FILING DATE: 1998-08-21

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 191

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence:/No. 6685944e =

; OTHER INFORMATION: synthetic construct

US-09-763-260-2

Query Match 100.0%; Score 1040; DB 4; Length 191;

Best Local Similarity 100.0%; Pred. No. 1.3e-92;

Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRPPQDVKFGGGQIVGGYLLPRGPRLGVRATKTSERSQPRG 60

Db 1 MSTNPKQKKNKNTNRPPQDVKFGGGQIVGGYLLPRGPRLGVRATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRITWAQPGYPMPLVNECGWAGLLSPRGSRPSWGPTDPRRRSRNLG 120

Db 61 RRQIPKARRPEGRITWAQPGYPMPLVNECGWAGLLSPRGSRPSWGPTDPRRRSRNLG 120

Qy 121 KVIDTLTCGFDLGMGYIPAVGAPLGGAAARALAHGVRVLEGVNATGNLPGCSFSIFLLA 180

Db 121 KVIDTLTCGFDLGMGYIPAVGAPLGGAAARALAHGVRVLEGVNATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191

Db 181 LLSCLTVPASA 191





PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,368  
; FILING DATE: 07/1991  
; FILING DATE: 13-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0205.001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2708  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3011 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-440-542-36

Query Match 99.5%; Score 1035; DB 1; Length 3011;  
Best Local Similarity 99.5%; Pred. No. 1.1e-90;  
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MSTNPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
DB 1 MSTNPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
QY 61 RRQPIPKARRPEGRGTWAQPGYPMPLYGNEGCGWAGWLLSPRGRSPSGPTDPRRRSRNLG 120  
DB 61 RRQPIPKARRPEGRGTWAQPGYPMPLYGNEGCGWAGWLLSPRGRSPSGPTDPRRRSRNLG 120  
QY 121 KVIDTLTCGADLMGYIPAVGAPLGGAARALAHGVRLDGVNYATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGADLMGYIPAVGAPLGGAARALAHGVRLDGVNYATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTVPASA 191  
DB 181 LLSCLTVPASA 191

## RESULT 5

US-07-910-760-10  
; Sequence 10, Application US/07910760  
; Patent No. 5683864  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-Lim  
; APPLICANT: Kuo, George  
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: P.O. Box 8097 (Int. Prop. R-440)  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/910,760  
; FILING DATE: 07-JUL-1992  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Blackburn Esq., Robert P.  
; REGISTRATION NUMBER: 30,447  
; REFERENCE/DOCKET NUMBER: 0101.002

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2702  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3011 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-910-760-10

Query Match 99.5%; Score 1035; DB 1; Length 3011;  
Best Local Similarity 99.5%; Pred. No. 1.1e-90;  
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MSTNPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
DB 1 MSTNPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRRGLGVRAATKTSERSQPRG 60  
QY 61 RRQPIPKARRPEGRGTWAQPGYPMPLYGNEGCGWAGWLLSPRGRSPSGPTDPRRRSRNLG 120  
DB 61 RRQPIPKARRPEGRGTWAQPGYPMPLYGNEGCGWAGWLLSPRGRSPSGPTDPRRRSRNLG 120  
QY 121 KVIDTLTCGADLMGYIPAVGAPLGGAARALAHGVRLDGVNYATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGADLMGYIPAVGAPLGGAARALAHGVRLDGVNYATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTVPASA 191  
DB 181 LLSCLTVPASA 191

## RESULT 6

US-08-440-519-10  
; Sequence 10, Application US/08440519  
; Patent No. 5712087  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-Lim  
; APPLICANT: Kuo, George  
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: P.O. Box 8097 (Int. Prop. R-440)  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/440,519  
; FILING DATE: 12-MAY-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/910,760  
; FILING DATE: 07-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Blackburn Esq., Robert P.  
; REGISTRATION NUMBER: 30,447  
; REFERENCE/DOCKET NUMBER: 0101.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2702  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3011 amino acids  
; TYPE: amino acid

QY	1	MSTNPKPKQKKNKNTNRRPQDVKFPGGGOIVGGVYLLPRRGPRLGVRATKTKTSERSQPRG	60
QY	1	MSTNPKPKQKKNKNTNRRPQDVKFPGGGOIVGGVYLLPRRGPRLGVRATKTKTSERSQPRG	60
Db	1	MSTNPKPKQKKNKNTNRRPQDVKFPGGGOIVGGVYLLPRRGPRLGVRATKTKTSERSQPRG	60
QY	61	RRQPIKARRPEGRITWAQPGYWPWLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120
Db	61	RRQPIKARRPEGRITWAQPGYWPWLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120
QY	121	KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRLDGVNATGNLPGCSFSIFLLA	180
Db	121	KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRLDGVNATGNLPGCSFSIFLLA	180
QY	181	LLSCLTVPASA 191	
Db	181	LLSCLTVPASA 191	
<p>RESULT 8</p> <p>US-08-440-210-36</p> <p>; Sequence 36, Application US/08440210</p> <p>; Patent No. 5766845</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Weiner, Amy J.</p> <p>; APPLICANT: Houghton, Michael</p> <p>; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions</p> <p>; NUMBER OF SEQUENCES: 45</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: Chiron Corporation</p> <p>; STREET: 4560 Horton Street</p> <p>; CITY: Emeryville</p> <p>; STATE: CA</p> <p>; COUNTRY: USA</p> <p>; ZIP: 94608</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Floppy disk</p> <p>; COMPUTER: IBM PC compatible</p> <p>; OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>; SOFTWARE: Patent In Release #1.0, Version #1.25</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/08/440,210</p> <p>; FILING DATE: 12-MAY-1995</p> <p>; CLASSIFICATION: 435</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/08/231,368</p> <p>; FILING DATE:</p> <p>; APPLICATION NUMBER: US 07/759,575</p> <p>; FILING DATE: 13-SEP-1991</p> <p>; ATTORNEY/AGENT INFORMATION:</p> <p>; NAME: McClung, Barbara G.</p> <p>; REGISTRATION NUMBER: 33,113</p> <p>; REFERENCE/DOCKET NUMBER: 0205.001</p> <p>; TELECOMMUNICATION INFORMATION:</p> <p>; TELEPHONE: (510) 601-2708</p> <p>; TELEFAX: (510) 655-3542</p> <p>; INFORMATION FOR SEQ ID NO: 36:</p> <p>; SEQUENCE CHARACTERISTICS:</p> <p>; LENGTH: 3011 amino acids</p> <p>; TYPE: amino acid</p> <p>; STRANDEDNESS: single</p> <p>; TOPOLOGY: linear</p> <p>; MOLECULE TYPE: protein</p> <p>US-08-440-210-36</p> <p>Query Match 99.5%; Score 1035; DB 1; Length 3011;</p> <p>Best Local Similarity 99.5%; Pred. No. 1.1e-90;</p> <p>Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</p>			
QY	1	MSTNPKPKQKKNKNTNRRPQDVKFPGGGOIVGGVYLLPRRGPRLGVRATKTKTSERSQPRG	60
Db	1	MSTNPKPKQKKNKNTNRRPQDVKFPGGGOIVGGVYLLPRRGPRLGVRATKTKTSERSQPRG	60
QY	61	RRQPIKARRPEGRITWAQPGYWPWLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120
Db	61	RRQPIKARRPEGRITWAQPGYWPWLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLG	120
QY	121	KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRLDGVNATGNLPGCSFSIFLLA	180
Db	121	KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRLDGVNATGNLPGCSFSIFLLA	180
QY	181	LLSCLTVPASA 191	
Db	181	LLSCLTVPASA 191	
<p>RESULT 7</p> <p>US-08-231-368-36</p> <p>; Sequence 36, Application US/08231368</p> <p>; Patent No. 5756312</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: Weiner, Amy J.</p> <p>; APPLICANT: Houghton, Michael</p> <p>; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions</p> <p>; NUMBER OF SEQUENCES: 45</p> <p>; CORRESPONDENCE ADDRESS:</p> <p>; ADDRESSEE: Chiron Corporation</p> <p>; STREET: 4560 Horton Street</p> <p>; CITY: Emeryville</p> <p>; STATE: CA</p> <p>; COUNTRY: USA</p> <p>; ZIP: 94608</p> <p>; COMPUTER READABLE FORM:</p> <p>; MEDIUM TYPE: Floppy disk</p> <p>; COMPUTER: IBM PC compatible</p> <p>; OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>; SOFTWARE: Patent In Release #1.0, Version #1.25</p> <p>; CURRENT APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US/08/231,368</p> <p>; FILING DATE:</p> <p>; CLASSIFICATION: 435</p> <p>; PRIOR APPLICATION DATA:</p> <p>; APPLICATION NUMBER: US 07/759,575</p> <p>; FILING DATE: 13-SEP-1991</p> <p>; ATTORNEY/AGENT INFORMATION:</p> <p>; NAME: McClung, Barbara G.</p> <p>; REGISTRATION NUMBER: 33,113</p> <p>; REFERENCE/DOCKET NUMBER: 0205.001</p> <p>; TELECOMMUNICATION INFORMATION:</p> <p>; TELEPHONE: (510) 601-2708</p> <p>; TELEFAX: (510) 655-3542</p> <p>; INFORMATION FOR SEQ ID NO: 36:</p> <p>; SEQUENCE CHARACTERISTICS:</p> <p>; LENGTH: 3011 amino acids</p> <p>; TYPE: amino acid</p> <p>; STRANDEDNESS: single</p> <p>; TOPOLOGY: linear</p> <p>; MOLECULE TYPE: protein</p> <p>US-08-231-368-36</p> <p>Query Match 99.5%; Score 1035; DB 1; Length 3011;</p> <p>Best Local Similarity 99.5%; Pred. No. 1.1e-90;</p> <p>Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;</p>			

Db 61 RRQPIKARRPEGRWTAQPGYPMPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPVGLGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

RESULT 9  
US-09-388-874-2  
; Sequence 2, Application US/09388874  
; Patent No. 6284249  
; GENERAL INFORMATION:  
; APPLICANT: Veronique Barban  
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR  
; FILE OF INVENTION: TREATING C HEPATITIS  
; FILE REFERENCE: PMCF97-03A  
; CURRENT APPLICATION NUMBER: US/09/388,874  
; EARLIER FILING DATE: 1999-09-02  
; EARLIER APPLICATION NUMBER: PCT/FR98/00448  
; EARLIER FILING DATE: 1998-03-06  
; EARLIER APPLICATION NUMBER: 97/02,887  
; EARLIER FILING DATE: 1997-03-06  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FASTSEQ for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 3011  
; TYPE: PRT  
; ORGANISM: Virus  
US-09-388-874-2

Query Match 99.5%; Score 1035; DB 3; Length 3011;  
Best Local Similarity 99.5%; Pred. No. 1.1e-90;  
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MSTNPKQKKNKNTNRPPQDVKPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
Db 1 MSTNPKQKKNKNTNRPPQDVKPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
Qy 61 RRQPIKARRPEGRWTAQPGYPMPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
Db 61 RRQPIKARRPEGRWTAQPGYPMPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPVGLGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

RESULT 10  
US-09-046-604-36  
; Sequence 36, Application US/09046604  
; Patent No. 6303292  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, Amy J.  
; APPLICANT: Houghton, Michael  
; TITLE OF INVENTION: Immunoreactive Polypeptide Compositions  
; NUMBER OF SEQUENCES: 45  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94608  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/046,604  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/231,368  
; FILING DATE:  
; APPLICATION NUMBER: US 07/759,575  
; FILING DATE: 13-SEP-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: McClung, Barbara G.  
; REGISTRATION NUMBER: 33,113  
; REFERENCE/DOCKET NUMBER: 0205.001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 601-2708  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 36:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3011 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-046-604-36  
Query Match 99.5%; Score 1035; DB 3; Length 3011;  
Best Local Similarity 99.5%; Pred. No. 1.1e-90;  
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MSTNPKQKKNKNTNRPPQDVKPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
Db 1 MSTNPKQKKNKNTNRPPQDVKPGGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
Qy 61 RRQPIKARRPEGRWTAQPGYPMPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
Db 61 RRQPIKARRPEGRWTAQPGYPMPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRSRNLG 120  
Qy 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPVGLGAPLGGAARALAHGVRVLEGGVNYATGNLPGCSFSIFLLA 180  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

RESULT 11  
US-08-440-549-10  
; Sequence 10, Application US/08440549  
; Patent No. 6312889  
; GENERAL INFORMATION:  
; APPLICANT: Houghton, Michael  
; APPLICANT: Choo, Qui-Lim  
; APPLICANT: Kuo, George  
; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)  
; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: P.O. Box 8097 (Int. Prop. R-440)  
; CITY: Emeryville  
; STATE: CA  
; COUNTRY: U.S.A.  
; ZIP: 94662-8097  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

```
; APPLICATION NUMBER: US/08/440,549
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/910,760
; FILING DATE: 07-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Blackburn Esq., Robert P.
; REGISTRATION NUMBER: 30,447
; REFERENCE/DOCKET NUMBER: 0101.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 601-2702
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3011 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-440-549-10

Query Match 99.5%; Score 1035; DB 3; Length 3011;
Best Local Similarity 99.5%; Pred. No. 1.1e-90;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPKKKKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
DB 1 MSTNPKPKKKKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
QY 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
DB 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVAGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191

RESULT 12
US-08-850-328-1
; Sequence 1, Application US/08850328
; Patent No. 6379886
; GENERAL INFORMATION:
; APPLICANT: TAKAHAMA, Y.
; APPLICANT: SHIRAI, J.
; TITLE OF INVENTION: DIAGNOSTIC REAGENT FOR HEPATITIS
; TITLE OF INVENTION: C VIRUS INFECTION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,328
; FILING DATE: 02-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mays, Thomas D

Query Match 99.5%; Score 1035; DB 3; Length 3011;
Best Local Similarity 99.5%; Pred. No. 1.1e-90;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPKKKKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
DB 1 MSTNPKPKKKKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
QY 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
DB 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVAGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191

RESULT 13
US-09-916-359-2
; Sequence 2, Application US/09916359
; Patent No. 6538123
; GENERAL INFORMATION:
; APPLICANT: Veronique Barban
; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
; TITLE OF INVENTION: TREATING C HEPATITIS
; FILE REFERENCE: PMCF97-03A
; CURRENT APPLICATION NUMBER: US/09/916,359
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 09/388,874
; PRIOR FILING DATE: 1999-09-02
; PRIOR APPLICATION NUMBER: 97/02,887
; PRIOR FILING DATE: 1997-03-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 3011
; TYPE: PRT
; ORGANISM: Virus
; US-09-916-359-2

Query Match 99.5%; Score 1035; DB 4; Length 3011;
Best Local Similarity 99.5%; Pred. No. 1.1e-90;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTNPKPKKKKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
DB 1 MSTNPKPKKKKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
QY 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
DB 61 RRQIPKARRPEGRRTWAQPGYWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRRRNLG 120
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYIPVAGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191
```

Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNATGNLPGCSFSIFLLA 180  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

RESULT 14  
US-07-681-703B-12  
; Sequence 12, Application US/07681703B  
; Patent No. 5443965  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory  
; APPLICANT: Kim, Jungsu P.  
; APPLICANT: Moeckli, Randolph  
; TITLE OF INVENTION: Hepatitis C Virus Epitopes  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave., Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/681,703B  
; FILING DATE: 05-APR-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 505,611  
; FILING DATE: 06-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 594,854  
; FILING DATE: 09-OCT-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 4600-076.21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 215 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-681-703B-12

Query Match 99.1%; Score 1031; DB 1; Length 215;  
Best Local Similarity 99.0%; Pred. No. 1.1e-91;  
Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKPKKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
Db 1 MGTNPKPKKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRGTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSMGPTDPRRSRNLG 120  
Db 61 RRQIPKARRPEGRGTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSMGPTDPRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

Search completed: August 19, 2005, 12:44:52  
Job time : 21 secs

RESULT 15  
US-07-681-703B-14  
; Sequence 14, Application US/07681703B  
; Patent No. 5443965  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory  
; APPLICANT: Kim, Jungsu P.  
; APPLICANT: Moeckli, Randolph  
; TITLE OF INVENTION: Hepatitis C Virus Epitopes  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Ave., Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/681,703B  
; FILING DATE: 05-APR-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 505,611  
; FILING DATE: 06-APR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 594,854  
; FILING DATE: 09-OCT-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fabian, Gary R.  
; REGISTRATION NUMBER: 33,875  
; REFERENCE/DOCKET NUMBER: 4600-076.21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 215 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-681-703B-14

Query Match 99.1%; Score 1031; DB 1; Length 215;  
Best Local Similarity 99.0%; Pred. No. 1.1e-91;  
Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MSTNPKPKKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
Db 1 MGTNPKPKKNTNRRPQDVKFGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60

Qy 61 RRQIPKARRPEGRGTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSMGPTDPRRSRNLG 120  
Db 61 RRQIPKARRPEGRGTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSMGPTDPRRSRNLG 120

Qy 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNATGNLPGCSFSIFLLA 180  
Db 121 KVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNATGNLPGCSFSIFLLA 180

Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191

Search completed: August 19, 2005, 12:44:52  
Job time : 21 secs

**This page blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 19, 2005, 12:28:45 ; Search time 117.5 Seconds  
(without alignments)  
628.691 Million cell updates/sec

Title: US-10-770-117-2

Perfect score: 1040

Sequence: 1 MSTNPKQKGNKNTNRPO.....CSPSIFLLALLSCLTVPASA 191

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1040	100.0	191	3	AAy82997 Hepatitis
2	1035	99.5	191	3	AAy82999 Hepatitis
3	1035	99.5	2816	2	AAr34009 HCV-1 pol
4	1035	99.5	2955	2	AAy14975 Amino aci
5	1035	99.5	2955	3	AAAB18541 Polypeptide
6	1035	99.5	3011	2	AAr21519 Compiled
7	1035	99.5	3011	2	AAr90931 Hepatitis
8	1035	99.5	3011	2	AAW34480 HCV polyp
9	1035	99.5	3011	2	AAW40038 HCV polyp
10	1035	99.5	3011	5	AAE22049 Hepatitis
11	1035	99.5	3011	8	ADL23107 Hepatitis
12	1035	99.5	3011	7	ADR29357 Hepatitis
13	1031	99.1	190	7	ADL39661 Hepatitis
14	1031	99.1	215	2	AAr14557 HCV Capsi
15	1031	99.1	215	2	AAr14558 Polypepti
16	1031	99.1	215	2	AAr81945 Hepatitis
17	1031	99.1	215	2	AAr81944 Hepatitis
18	1031	99.1	217	2	AAW70828 Hepatitis
19	1031	99.1	217	2	AAW70829 Insert of
20	1031	99.1	217	2	AAW80517 Insert of
21	1031	99.1	217	2	AAW80516 Hepatitis
22	1030	99.0	190	5	AAW71256 HCV type
23	1030	99.0	2955	8	ADN35978 HCV cDNA
24	1026	98.7	191	2	AAr44010 Hepatitis
25	1026	98.7	191	2	AAr92938 Hepatitis

#### ALIGNMENTS

RESULT 1  
AAy82997

ID AAY82997 standard; peptide; 191 AA.

XX AAY82997;

DT 04-JUL-2000 (first entry)

DE Hepatitis C virus core polypeptide having Leu to Ala substitution.

KW Hepatitis C virus; HCV; core polypeptide; antigen; epitope; agetope;

KW anchor residue; cytotoxic T lymphocyte; CTL; immune response; MHC;

KW Major histocompatibility complex; vaccine; treatment.

XX Synthetic.

OS Hepatitis C virus; (HCV).

XX Key Location/Qualifiers

FT Misc-difference 139 /label= Ala substituted for Leu

XX WO200011186-A1.

XX 02-MAR-2000.

XX 17-AUG-1999; 99WO-US018674.

XX 21-AUG-1998; 98US-0097446P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX Berzofsky JA, Sarobe P, Pendleton CD, Feinstein SM, Arichi T;

XX Major ME;

XX WPI; 2000-246569/21.

PT Hepatitis C virus polypeptides is useful as a vaccine for treating  
Hepatitis C virus infection and for activating cytotoxic T lymphocytes.

XX Claim 2; Page 71-72; 78pp; English.

XX Peptide fragments of the Hepatitis C virus (HCV) core polypeptide can be  
used to elicit an immune response. A cytotoxic T-lymphocyte (CTL)  
response is present in patients acutely and chronically infected by HCV.  
CTL's recognise viral antigens as peptides presented by Class I molecules  
of the major histocompatibility complex (MHC). These peptide antigens are  
typically 8-10 amino acids long. These peptide antigens typically have a  
sequence pattern or motif depending upon which particular MHC antigen

AAr92941 Hepatitis  
AAr92937 Hepatitis  
AAr92939 Hepatitis  
AAr92940 Hepatitis  
AAy94411 Human hep  
Adj10436 HCV Core  
Adj10438 HCV Core  
AAr79222 pHCV141-e  
AAw12715 HCV genom  
AAr79221 pHCV176-e  
AAe00449 Hepatitis  
AAe00447 Hepatitis  
AAe00442 Hepatitis  
Adm4822 Hepatitis  
AAr40119 HCV genom  
AAr40120 HCV genom  
AAr79232 HCV seque  
AAw77397 Hepatitis  
AAw77398 Hepatitis  
AAw98020 Infectiou

26 1026 98.7 191 2 AAR92941  
27 1026 98.7 191 2 AAR92937  
28 1026 98.7 191 2 AAR92939  
29 1026 98.7 191 2 AAR92940  
30 1026 98.7 191 3 AAY94411  
31 1026 98.7 249 8 ADJ10436  
32 1026 98.7 473 8 ADJ10438  
33 1026 98.7 967 2 AAR79222  
34 1026 98.7 1006 2 AAW12715  
35 1026 98.7 1648 2 AAR79221  
36 1026 98.7 2984 4 AAe00449  
37 1026 98.7 2984 4 AAe00447  
38 1026 98.7 2984 4 AAe00442  
39 1026 98.7 3002 7 ADM4822  
40 1026 98.7 3011 2 AAR40119  
41 1026 98.7 3011 2 AAR40120  
42 1026 98.7 3011 2 AAR79232  
43 1026 98.7 3011 2 AAW77397  
44 1026 98.7 3011 2 AAW77398  
45 1026 98.7 3011 2 AAW98020

CC they are bound and presented by. The motifs are based on the presence in  
CC precise positions in the peptide sequence of several amino acids  
CC (agretropic residues) called anchor residues, responsible for interactions  
CC between peptide and MHC molecule, as well as other secondary position  
CC that may help to stabilise the interaction. Thus, single amino acid  
CC changes within the peptide sequence can enhance the immune response.  
CC Peptides derived from the HCV core polypeptide having substituted amino  
CC acids within their sequence, specifically a hepatitis C virus core  
CC polypeptide comprising an L to an A substitution at amino acid position  
CC 139 enhance the immune response against HCV. They can therefore be used  
CC as a vaccine or for treating HCV infection  
XX  
XX Sequence 191 AA;

Query Match 100.0%; Score 1040; DB 3; Length 191;  
Best Local Similarity 100.0%; Pred. No. 2.7e-91;  
Matches 191; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSTNPKQKKNKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60  
DB 1 MSTNPKQKKNKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60  
QY 61 RRQPIPKARRPEGRWTWAPQGYWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120  
DB 61 RRQPIPKARRPEGRWTWAPQGYWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120  
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTVPASA 191  
DB 181 LLSCLTVPASA 191

RESULT 2  
AAY82999  
ID AAY82999 standard; peptide; 191 AA.  
XX AC AAY82999;  
XX  
DT 12-SEP-2003 (revised)  
DT 04-JUL-2000 (first entry)  
XX Hepatitis C virus core polypeptide.  
XX  
XX Hepatitis C virus; HCV; core polypeptide; antigen; epitope; agretope;  
XX anchor residue; cytotoxic T lymphocyte; CTL; immune response; MHC;  
XX Major histocompatibility complex; vaccine; treatment.  
XX  
XX Hepatitis C virus; (HCV).  
XX WO200011186-A1.  
XX  
XX 02-MAR-2000.  
XX 17-AUG-1999; 99WO-US018674.  
XX  
XX 21-AUG-1998; 98US-0097446P.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Berzofsky JA, Sarobe P, Pendleton CD, Feinstein SM, Arichi T;  
XX Major ME;  
XX WPI; 2000-246569/21.  
XX  
XX Hepatitis C virus polypeptides is useful as a vaccine for treating  
XX Hepatitis C virus infection and for activating cytotoxic T lymphocytes.  
XX  
XX Disclosure; Page 72-73; 78pp; English.  
XX  
XX Peptide fragments of the Hepatitis C virus (HCV) core polypeptide can be

CC used to elicit an immune response. A cytotoxic T-lymphocyte (CTL)  
CC response is present in patients acutely and chronically infected by HCV.  
CC CTL's recognise viral antigens as peptides presented by Class I molecules  
CC of the major histocompatibility complex (MHC). These peptide antigens are  
CC typically 8-10 amino acids long. These peptide antigens typically have a  
CC sequence pattern or motif depending upon which particular MHC antigen  
CC they are bound and presented by. The motifs are based on the presence in  
CC precise positions in the peptide sequence of several amino acids  
CC (agretropic residues) called anchor residues, responsible for interactions  
CC between peptide and MHC molecule, as well as other secondary position  
CC that may help to stabilise the interaction. Thus, single amino acid  
CC changes within the peptide sequence can enhance the immune response.  
CC Peptides derived from the HCV core polypeptide having substituted amino  
CC acids within their sequence, specifically a hepatitis C virus core  
CC polypeptide comprising an L to an A substitution at amino acid position  
CC 139 enhance the immune response against HCV. They can therefore be used  
CC as a vaccine or for treating HCV infection. (Updated on 12-SEP-2003 to  
CC standardise OS field)  
XX  
XX Sequence 191 AA;

Query Match 99.5%; Score 1035; DB 3; Length 191;  
Best Local Similarity 99.5%; Pred. No. 8.1e-91;  
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MSTNPKQKKNKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60  
DB 1 MSTNPKQKKNKNTNRRPDVKFPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60  
QY 61 RRQPIPKARRPEGRWTWAPQGYWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120  
DB 61 RRQPIPKARRPEGRWTWAPQGYWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRNLG 120  
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPAVGAPLGGAAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTVPASA 191  
DB 181 LLSCLTVPASA 191

RESULT 3  
AAR34009  
ID AAR34009 standard; protein; 2816 AA.  
XX AC AAR34009;  
XX  
DT 25-MAR-2003 (revised)  
DT 26-JUL-1993 (first entry)  
XX  
XX HCV-1 polyprotein.  
XX  
XX Polymerase chain reaction; PCR; amplify; primer; hepatitis C virus; HCV;  
XX asymptomatic; chronically infected; epitope; viral isolate; domain;  
XX immunological; cross-reactive; envelope protein; vaccine;  
XX gp53 (BVDV)/gp55; hog cholera virus; pestivirus; NS1; flavivirus.  
XX  
XX Hepatitis C virus.  
XX WO9306126-A1.  
XX  
XX 01-APR-1993.  
XX 11-SEP-1992; 92WO-US007683.  
XX 13-SEP-1991; 91US-00759575.  
XX (CHIR ) CHIRON CORP.  
XX  
XX Weiner AJ, Houghton M;  
XX WPI; 1993-117468/14.  
XX



XX Immuno-reactive hepatitis C virus polypeptide compsns. - contg. at least  
PT 2 sequences from the first variable domain of distinct HCV isolates.  
XX  
XX Disclosure; Fig 9; 106pp; English.  
XX  
XX This sequence represents the entire hepatitis C virus polypeptide. HCV is  
CC a member of the flavivirus family and appears to encode a basic  
CC polypeptide domain ("C") at the N-terminal of the viral polypeptide,  
CC followed by two glycoprotein domains ("E1", "E2/NS1"), upstream of the  
CC nonstructural genes NS2 through NS5. See also AAQ39134-48, AAQ33982- 4008  
CC and AAQ38088-89. (Updated on 25-MAR-2003 to correct PN field.)  
XX  
SQ Sequence 2816 AA;  
  
Query Match 99.5%; Score 1035; DB 2; Length 2816;  
Best Local Similarity 99.5%; Pred. No. 1.8e-89;  
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
DB 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
61 RRQPIKARRPEGRTPAQPGYPWPLYGNEGCGWAGLLSPRGSRPSWGPTDPRRRSRNLG 120  
61 RRQPIKARRPEGRTPAQPGYPWPLYGNEGCGWAGLLSPRGSRPSWGPTDPRRRSRNLG 120  
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGVNATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGVNATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTVPASA 191  
DB 181 LLSCLTVPASA 191  
  
RESULT 4  
AA14975  
ID AA14975 standard; protein; 2955 AA.  
XX  
XX AA14975;  
XX  
XX 20-MAR-2003 (revised)  
DT 08-NOV-1999 (first entry)  
XX  
XX Amino acid sequence of HCV-1 ORF.  
XX  
XX Hepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH;  
KW HCV infection; vaccine.  
XX  
XX Hepatitis C virus.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 441  
FT /note= "encoded by TT"  
FT Misc-difference 461  
FT /note= "encoded by CCCC"  
XX  
XX EP939128-A2.  
XX  
XX 01-SEP-1999.  
XX  
XX 17-SEP-1990; 99EP-00101746.  
XX  
XX 15-SEP-1989; 89US-00408045.  
PR 21-DEC-1989; 89US-00456142.  
PR 17-SEP-1990; 90EP-00310149.  
XX  
XX (OYAA/) OYA A.  
PA (CHIR ) CHIRON CORP.  
XX  
XX Miyamura T, Saito I, Houghton M, Weiner AJ, Han J, Kolberg JA;  
PI Cha T, Irvine BD;

XX WPI; 1999-480843/41.  
DR N-PSDB; AAQ07656.  
XX  
XX New Hepatitis C Virus isolates, useful for diagnosis of hepatitis  
PT infections and development of vaccines.  
XX  
XX Disclosure; Fig 12; 132pp; English.  
XX  
XX The invention provides two new isolates of hepatitis C virus (HCV), J1  
CC and J7. These two isolates comprise nucleotide and amino acid sequences  
CC that are distinct from the HCV isolate HCV-1. The nucleotide sequences  
CC may be used to detect non-A, non-B HCV (NANBH) polynucleotides by  
CC hybridization for diagnosis of NANBH infections. They may also be used to  
CC screen blood donors, donated blood and blood products for this infection.  
CC The isolates may also be used to isolate other naturally occurring  
CC variants of the virus. The polypeptides may be used as a vaccine for  
CC administration to patients to protect against infection with NANBH. The  
CC present sequence represents the amino acid sequence of HCV-1 ORF.  
CC (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to  
CC correct PR field.)  
XX  
SQ Sequence 2955 AA;  
  
Query Match 99.5%; Score 1035; DB 2; Length 2955;  
Best Local Similarity 99.5%; Pred. No. 1.9e-89;  
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
DB 1 MSTNPKQKKNKNTNRRPQDVKPGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
61 RRQPIKARRPEGRTPAQPGYPWPLYGNEGCGWAGLLSPRGSRPSWGPTDPRRRSRNLG 120  
61 RRQPIKARRPEGRTPAQPGYPWPLYGNEGCGWAGLLSPRGSRPSWGPTDPRRRSRNLG 120  
QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGVNATGNLPGCSFSIFLLA 180  
DB 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEGVNATGNLPGCSFSIFLLA 180  
QY 181 LLSCLTVPASA 191  
DB 181 LLSCLTVPASA 191  
  
RESULT 5  
AA18541  
ID AA18541 standard; protein; 2955 AA.  
XX  
XX AA18541;  
XX  
XX 15-JAN-2001 (first entry)  
XX  
XX Polypeptide encoded by sense strand of HCV.  
XX  
XX Hepatitis C virus; HCV; antisense polynucleotide; polypeptide;  
KW viral infectivity; viral replication.  
XX  
XX Hepatitis C virus.  
XX  
XX EP1034785-A2.  
XX  
XX 13-SEP-2000.  
XX  
XX 16-MAR-1990; 2000EP-00109602.  
XX  
XX 17-MAR-1989; 89US-00325338.  
PR 20-APR-1989; 89US-00341334.  
PR 18-MAY-1989; 89US-00355002.  
PR 16-MAR-1990; 90EP-00302866.  
XX  
XX (CHIR ) CHIRON CORP.  
XX  
XX

```

PI Houghton M, Choo Q, Kuo G;
XX WPI; 2000-566891/53.
DR N-PSDB; AAA75297.
XX
XX Novel composition comprising a hepatitis C virus antisense polynucleotide
XX which is complementary to or corresponds to a sense strand of the virus
XX genome, and selectively hybridizes to it.
XX
XX Example; Fig 17; 75pp; English.
XX
XX The specification describes a pharmaceutical composition which comprises
XX a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
XX characterized by a positive stranded RNA genome which has 40% homology at
XX the polypeptide level to a HCV polyprotein. The antisense polynucleotide
XX binds to cellular polynucleotides which enhance and/or are required for
XX viral infectivity, replicative ability or chronicity. The antisense
XX polynucleotides may also be designed to bind with high specificity, to be
XX of increased stability, to be stable and to have low toxicity. The
XX composition also comprises an agent which causes viral RNA to be
XX inactive. The composition is used for preventing HCV replication in a
XX system. The present sequence is encoded by a novel HCV cDNA sequence,
XX which is used in the course of the invention
XX
XX Sequence 2955 AA;
SQ
Query Match 99.5%; Score 1035; DB 3; Length 2955;
Best Local Similarity 99.5%; Pred. No. 1.9e-89;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MSTNPKPKKKNTNRNRRPDDVKEPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
DB 1 MSTNPKPKKKNTNRNRRPDDVKEPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
QY 61 RRQPIKARPEGTWAPQGYPPWPLYNCGGAGWALLSPRSGRPSWGTDPDRRRSRNLG 120
DB 61 RRQPIKARPEGTWAPQGYPPWPLYNCGGAGWALLSPRSGRPSWGTDPDRRRSRNLG 120
QY 121 KVIDTLTCGFADLMGYITPAVGAPLGGAAARALAHGVRVLEDGVNATGNLPGCSFSIFLLA 180
DB 121 KVIDTLTCGFADLMGYITPAVGAPLGGAAARALAHGVRVLEDGVNATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
DB 181 LLSCLTVPASA 191
RESULT 6
AAR21519
ID AAR21519 standard; protein; 3011 AA.
XX
XX AAR21519;
AC
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 22-JUN-1992 (first entry)
XX
XX Compiled HCV sequence.
DE
XX HCV1; serum; gt11.
XX
XX Hepatitis C virus type 1.
OS
XX Key Location/Qualifiers
FH Misc-difference 9 /label= ARG
FT Misc-difference 11 /label= THR
FT Misc-difference 176 /label= THR
FT Misc-difference 334 /label= VAL
FT Misc-difference 603

```

```

FT Misc-difference 848 /label= ILE
FT Misc-difference 1114 /label= (ASN)
FT Misc-difference 1114 /label= SER
FT Misc-difference 1117 /label= THR
FT Misc-difference 1276 /label= LEU
FT Misc-difference 1328 /label= (VAL)
FT Misc-difference 1454 /label= TYR
FT Misc-difference 1471 /label= (SER)
FT Misc-difference 1877 /label= (GLY)
FT Misc-difference 1948 /label= (HIS)
FT Misc-difference 1949 /label= (CYS)
FT Misc-difference 2021 /label= (VAL)
FT Misc-difference 2349 /label= (SER)
FT Misc-difference 2385 /label= (PHE)
FT Misc-difference 2386 /label= (ALA)
FT Misc-difference 2502 /label= (PHE)
FT Misc-difference 2690 /label= (GLY)
FT Misc-difference 2996 /label= (PRO)
XX
XX W09202642-A.
XX
XX 20-FEB-1992.
XX
XX 10-AUG-1990; 90US-00566209.
XX
XX 10-AUG-1990; 90US-00566209.
XX (CHIR ) CHIRON CORP.
XX
XX Houghton M, Choo QL, Kuo G, Weiner AJ, Uredea MS, Irvine BD;
XX Kolberg JA;
XX WPI; 1992-080094/10.
XX N-PSDB; AAQ211744.
XX
XX Reagents for isolating, amplifying and detecting HCV polynucleotide(s) -
XX used to monitor spread of blood-borne non-a, non-b hepatitis virus
XX infection and screen blood samples for virus.
XX
XX Disclosure; Fig 1; 67pp; English.
XX
XX Heterogeneities in cloned DNAs of HCV1 are indicated by the amino acid
XX indicated in the features, the parentheses indicated that the
XX heterogeneity was detected at or near to the 5'- or 3'-end of the HCV in
XX the clone. The sequence is derived from a composite HCV cDNA from HCV1, a
XX prototypic HCV. The DNA sequence is based upon sequence information
XX derived from a no. of HCV cDNA clones, which were isolated from a no. of
XX HCV cDNA libraries, including the "c" library present in lambda gt11
XX (ATCC No.40394), and from human serum. The HCV cDNA clones were isolated
XX by methods described in WO9014436. The clones from which the sequence was
XX derived are 5'clone32, b114a, 18g, ag30a, CA205a, CA290a, CA216a, p114a,
XX CA167b, CA156e, CA84a, CA59a, K9-1 (also called k9-1), 26j, 13i, 12f,
XX 14i, 11b, 7f, 7e, 8h, 33c, 40b, 37b, 35, 36, 8i, 32, 33b, 25c, 14c, 8f,
XX 33f, 33g, 33c, 35f, 19g, 26g, 15e, b5a, 16jh, 6k and p131jh. (Updated on
XX 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI
XX field.) (Updated on 24-OCT-2003 to standardise OS field)

```

```
XX SQ Sequence 3011 AA;
Query Match 99.5%; Score 1035; DB 2; Length 3011;
Best Local Similarity 99.5%; Pred. No. 1.9e-89;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Qy 61 RRQPIPKARRPEGRTPWPGYPWPLYGNECGWAGWLLSPRGSRPSPGPTDPRRRSRNLG 120
Db 61 RRQPIPKARRPEGRTPWPGYPWPLYGNECGWAGWLLSPRGSRPSPGPTDPRRRSRNLG 120
Qy 121 KVDTLTGCFADLMGYIPAVGAPLGGAARALAHGVRVLEGVNATGNLPGCSFSIFLLA 180
Db 121 KVDTLTGCFADLMGYIPAVGAPLGGAARALAHGVRVLEGVNATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 7
ID AAR90931 standard; protein; 3011 AA.
XX AAR90931;
AC AC
XX 25-MAR-2003 (revised)
DT 15-MAY-1996 (first entry)
XX Hepatitis C virus polypeptide.
DE Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;
KW antibodies.
XX Hepatitis C virus.
OS
XX Key Location/Qualifiers
FH Misc-difference 1.122
FT /label= antigen
FT /note= "C22; AAR90936"
FT Misc-difference 199.328
FT /label= antigen
FT /note= "S2; AAR90935"
FT Misc-difference 1192.1457
FT /label= antigen
FT /note= "C33C; AAR90932"
FT Misc-difference 1569.1931
FT /label= antigen
FT /note= "C100; AAR90933"
FT Misc-difference 2054.2464
FT /label= antigen
FT /note= "NS5; AAR90934"
XX EP693687-A1.
XX 24-JAN-1996.
XX 03-APR-1991; 95EP-00114016.
XX 04-APR-1990; 90US-00504352.
XX (CHIR ) CHIRON CORP.
XX Houghton M, Choo Q, Kuo G;
XX WPI; 1996-117956/13.
XX N-PSDB; AAT12710.
XX Combinations of synthetic Hepatitis C Virus antigens - provide more
```

```
PT effective diagnosis of Non-A, Non-B Hepatitis.
XX Disclosure; Fig 1(A-Y); 53pp; English.
XX The combination comprises an HCV antigen from the C domain (pref. C22 -
CC AAR90936) and at least one HCV antigen from the NS3 (pref. C33c -
CC AAR90932), NS4 (pref. C100 - AAR90933), S (pref. S2 - AAR90933) or NS5
CC (AAR90934) domain. The antigens may in the form of a fusion protein, a
CC simple physical mixture, or the individual antigens commonly bound to a
CC solid matrix. They are pref. prepd. by recombinant DNA techniques
CC (primers are given in AAT12711-T12716), but can be synthesised or
CC isolated from HCV using affinity chromatography. (Updated on 25-MAR-2003
CC to correct PF field.)
XX
SQ Sequence 3011 AA;
Query Match 99.5%; Score 1035; DB 2; Length 3011;
Best Local Similarity 99.5%; Pred. No. 1.9e-89;
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MSTNPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Db 1 MSTNPKQKKNKNTNRRPQDVKPGGGQIVGGVYLLPRGPRLGVRATKTSERSQPRG 60
Qy 61 RRQPIPKARRPEGRTPWPGYPWPLYGNECGWAGWLLSPRGSRPSPGPTDPRRRSRNLG 120
Db 61 RRQPIPKARRPEGRTPWPGYPWPLYGNECGWAGWLLSPRGSRPSPGPTDPRRRSRNLG 120
Qy 121 KVDTLTGCFADLMGYIPAVGAPLGGAARALAHGVRVLEGVNATGNLPGCSFSIFLLA 180
Db 121 KVDTLTGCFADLMGYIPAVGAPLGGAARALAHGVRVLEGVNATGNLPGCSFSIFLLA 180
Qy 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 8
AAW34480
ID AAW34480 standard; protein; 3011 AA.
XX AAW34480;
AC AC
XX 25-MAR-2003 (revised)
DT 16-MAR-1998 (first entry)
XX HCV polypeptide.
DE PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
XX C domain; S domain; NS5; HCV polypeptide; anti-HCV antibody; detection;
XX NS4.
XX Hepatitis C virus.
OS
XX Key Location/Qualifiers
FH Misc-difference 366
FT /note= "can optionally be Arg"
FT Misc-difference 372
FT /note= "can optionally be Thr"
FT Misc-difference 867
FT /note= "can optionally be Thr"
FT Misc-difference 1341
FT /note= "can optionally be Val"
FT Misc-difference 2148
FT /note= "can optionally be Ile"
FT Misc-difference 2883
FT /note= "can optionally be Asn"
FT Misc-difference 3681
FT /note= "can optionally be Ser"
FT Misc-difference 3690
FT /note= "can optionally be Thr"
FT Misc-difference 4167
FT /note= "can optionally be Leu"
```



FT also be Leu"  
FT 1328 Modified-site /note= "As given in the specification this amino acid can  
FT also be Val"  
FT 1452 Modified-site /note= "As given in the specification this amino acid can  
FT also be Tyr"  
FT 1472 Modified-site /note= "As given in the specification this amino acid can  
FT also be Ser"  
FT 1640 Domain /label= NS4\_domain  
FT 1877 Modified-site /note= "As given in the specification this amino acid can  
FT also be Gly"  
FT 1948 Modified-site /note= "As given in the specification this amino acid can  
FT also be His"  
FT 1949 Modified-site /note= "As given in the specification this amino acid can  
FT also be Cys"  
FT 2000 Domain /label= NS5\_domain  
FT 2021 Modified-site /note= "As given in the specification this amino acid can  
FT also be Val"  
FT 2348 Modified-site /note= "As given in the specification this amino acid can  
FT also be Ser"  
FT 2385 Modified-site /note= "As given in the specification this amino acid can  
FT also be Phe"  
FT 2386 Modified-site /note= "As given in the specification this amino acid can  
FT also be Ala"  
FT 2502 Modified-site /note= "As given in the specification this amino acid can  
FT also be Phe"  
FT 2690 Modified-site /note= "As given in the specification this amino acid can  
FT also be Gly"  
FT 2921 Modified-site /note= "As given in the specification this amino acid can  
FT also be Gly"  
FT 2996 Modified-site /note= "As given in the specification this amino acid can  
FT also be Pro"  
XX US5712087-A.  
PN  
XX  
XX 27-JAN-1998.  
PD  
XX  
XX 12-MAY-1995; 95US-00440519.  
XX  
XX 04-APR-1990; 90US-00504352.  
PR  
XX 07-JUL-1992; 92US-00910760.  
PR  
XX (CHIR ) CHIRON CORP.  
PA  
XX  
XX Kuo G, Houghton M, Choo Q;  
PI  
XX WPI, 1998-119973/11.  
DR  
XX N-PSDB; AAV09989.  
XX  
XX Immunassays for hepatitis C virus antibodies - using combinations of  
PT antigenic fragments of HCV polypeptide.  
XX  
XX Disclosure; Fig 1; 59pp; English.  
XX  
XX This sequence represents the hepatitis C virus (HCV) polypeptide which is  
CC used in the construction of novel combinations of HCV antigens that have  
CC a broader range of immunological activity than any single HCV antigen. An  
CC example of such an antigen given in this specification comprises a first

CC antigen containing at least 8 amino acids of the C domain of the HCV  
CC polypeptide and a second antigen comprising at least 8 amino acids of the  
CC NS3 domain, the NS4 domain, the S domain or the NS5 domain of the HCV  
CC polypeptide in the form of a fusion protein, a physical mixture or bound  
CC to a solid matrix. Note: The features given in the specification as  
CC represented in the feature table of AAW40038 differ from the positions  
CC indicated in Figure 1  
XX  
SQ Sequence 3011 AA;  
Query Match 99.5%; Score 1035; DB 2; Length 3011;  
Best Local Similarity 99.5%; Pred. No. 1.9e-89;  
Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MSTNPKQKKNKNTNRQDVKFGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
Db 1 MSTNPKQKKNKNTNRQDVKFGGQIVGGVYLLPRRGPRLGVRATKTSERSQPRG 60  
Qy 61 RRQPIPKARRPEGRWAPQGYPMPLYGNECGWAGWLLSPRGSRPSPGTPDPRRRRNLG 120  
Db 61 RRQPIPKARRPEGRWAPQGYPMPLYGNECGWAGWLLSPRGSRPSPGTPDPRRRRNLG 120  
Qy 121 KVDTTTCGFADLMGYIPVAGAPLGGAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180  
Db 121 KVDTTTCGFADLMGYIPVAGAPLGGAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180  
Qy 181 LLSCLTVPASA 191  
Db 181 LLSCLTVPASA 191  
RESULT 10  
AAE22049  
ID AAE22049 standard; protein; 3011 AA.  
XX  
AC AAE22049;  
XX  
XX 16-JUL-2002 (first entry)  
DT  
XX  
XX Hepatitis C virus (HCV) polypeptide.  
DE  
XX  
XX Hepatitis C virus; HCV; antigen; C domain; polypeptide; NS3 domain;  
KW NS4 domain; S domain; NS5 domain.  
XX  
XX Hepatitis C virus.  
OS  
XX  
XX Key Location/Qualifiers  
FH Domain 1. .122  
FT /label= C domain  
FT Domain 199. .328  
FT /label= S domain  
FT Region 1192. .1931  
FT /note= "c200 polypeptide"  
FT Region 1192. .1457  
FT /note= "NS3 domain antigen"  
FT Region 1569. .1931  
FT /note= "NS4 antigen"  
FT Region 2054. .2464  
FT /note= "NS5 antigen"  
XX  
XX US6312889-B1.  
XX  
XX  
XX 06-NOV-2001.  
PD  
XX  
XX 12-MAY-1995; 95US-00440549.  
XX  
XX 04-APR-1990; 90US-00504352.  
PR  
XX 07-JUL-1992; 92US-00910760.  
PR  
XX (CHIR ) CHIRON CORP.  
PA  
XX Houghton M, Choo Q, Kuo G;  
PI  
XX

DR WPI; 2002-040268/05.  
 XX N-PSDB; AAD35043.  
 PT Combination of hepatitis C viral (HCV) antigens, useful in improved  
 XX immunoassay for detecting HCV antibodies.  
 PS Example 1; Col 45-60; 58pp; English.  
 XX The invention relates to combination of hepatitis C viral (HCV) antigens  
 CC that have a broader range of immunological reactivity than any single HCV  
 CC antigen. The combinations consist of an antigen from the C domain of the  
 CC HCV polyprotein, and at least one additional HCV antigen from either the  
 CC NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in  
 CC the form of fusion protein, a simple physical mixture, or the individual  
 CC antigens commonly bound to a solid matrix. The combinations of antigens  
 CC provides broad range immunoassays for anti-HCV antibodies. The invention  
 CC therefore provides a method for detecting antibodies to HCV in a mammal  
 CC suspected of containing such antibodies. The present sequence is HCV  
 CC polyprotein. Note: This sequence SEQ.ID.NO:10 is stated to be similar to  
 CC the sequence shown in Fig 1 (AAB22052) of the specification. However  
 CC these sequences differ  
 XX  
 SQ Sequence 3011 AA;  
 Query Match 99.5%; Score 1035; DB 5; Length 3011;  
 Best Local Similarity 99.5%; Pred. No. 1.9e-89;  
 Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MSTNPKPKQKKNKNTNRRPQDVKFGGGQIVGGVYLLPRRGLGVRATKTSERSQPRG 60  
 DB 1 MSTNPKPKQKKNKNTNRRPQDVKFGGGQIVGGVYLLPRRGLGVRATKTSERSQPRG 60  
 QY 61 RRQPIPKARPEGRGRTWAQPGYPWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRLG 120  
 DB 61 RRQPIPKARPEGRGRTWAQPGYPWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRLG 120  
 QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPVPLVGAAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTVPASA 191  
 DB 181 LLSCLTVPASA 191  
 RESULT 11  
 ADL23107  
 ID ADL23107 standard; protein; 3011 AA.  
 XX  
 AC ADL23107;  
 XX  
 DT 20-MAY-2004 (first entry)  
 XX  
 DE Hepatitis C virus protein sequence Seqid 2.  
 XX  
 KW hepatitis C virus; HCV; viral envelope; E2 glycoprotein;  
 KW low density lipoprotein; LDL; HCV infection; hypercholesterolaemia;  
 KW hyperlipidaemia; coronary heart disease; hepatotropic; virucidal;  
 KW plasma lipoprotein.  
 XX  
 OS Hepatitis C virus.  
 XX  
 XX WO2004003141-A2.  
 PN  
 PD 08-JAN-2004.  
 XX  
 PP 24-JUN-2003; 2003WO-US019834.  
 XX  
 PR 28-JUN-2002; 2002US-0392158P.  
 PR 27-MAY-2003; 2003US-0044572A.  
 XX  
 PA (IOWA ) UNIV IOWA RES FOUND.  
 XX

PI Stapleton JT, Wuenschmann S;  
 XX  
 DR WPI; 2004-083029/08.  
 DR N-PSDB; ADL23106.  
 XX  
 PT Use of Hepatitis C Virus E2 glycoprotein in reducing low density  
 XX lipoprotein levels in a subject or in treating or preventing hepatitis C  
 PT virus infection.  
 PS Disclosure; SEQ ID NO 2; 135pp; English.  
 XX  
 CC This invention relates to a novel method for identifying inhibitors of  
 CC the hepatitis C virus (HCV), and also the use of an HCV viral envelope  
 CC protein identified as E2 glycoprotein in reducing low density lipoprotein  
 CC (LDL) levels. Specifically, it refers to the knowledge that HCV E2  
 CC glycoprotein binds to the lipid moiety of human lipoproteins and uses the  
 CC natural LDL receptor to bind to cells, thus identifies a novel route by  
 CC which HCV gains entry to the cell providing a novel therapeutic target.  
 CC The present invention describes screening methods to identify inhibitors  
 CC of HCV infection that comprises admixing a candidate substance with the  
 CC E2 glycoprotein and a plasma lipoprotein in order to determine a  
 CC reduction in binding in the presence of an appropriate inhibitor. As  
 CC such, these compositions can be used to reduce hypercholesterolaemia and  
 CC in turn treat hyperlipidaemia and associated conditions including  
 CC coronary heart disease. Accordingly, they exhibit hepatotropic and  
 CC virucidal activities. This polypeptide sequence is the HCV protein  
 CC sequence of the invention.  
 XX  
 SQ Sequence 3011 AA;  
 Query Match 99.5%; Score 1035; DB 8; Length 3011;  
 Best Local Similarity 99.5%; Pred. No. 1.9e-89;  
 Matches 190; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MSTNPKPKQKKNKNTNRRPQDVKFGGGQIVGGVYLLPRRGLGVRATKTSERSQPRG 60  
 DB 1 MSTNPKPKQKKNKNTNRRPQDVKFGGGQIVGGVYLLPRRGLGVRATKTSERSQPRG 60  
 QY 61 RRQPIPKARPEGRGRTWAQPGYPWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRLG 120  
 DB 61 RRQPIPKARPEGRGRTWAQPGYPWPPLYGNECGWAGWLLSPRGSRPSWGPTDPRRSRLG 120  
 QY 121 KVIDTLTCGFADLMGYIPAVGAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
 DB 121 KVIDTLTCGFADLMGYIPVPLVGAAPLGGAARALAHGVRVLEDDGVNYATGNLPGCSFSIFLLA 180  
 QY 181 LLSCLTVPASA 191  
 DB 181 LLSCLTVPASA 191  
 RESULT 12  
 ADL29357  
 ID ADL29357 standard; protein; 3011 AA.  
 XX  
 AC ADL29357;  
 XX  
 DT 18-NOV-2004 (first entry)  
 XX  
 DE Hepatitis C virus polyprotein precursor protein SEQ ID NO:1.  
 XX  
 KW anti-hepatitis C virus; anti-HCV; immune response; anti-HIV; virucide;  
 KW antinflammatory; hepatotropic; immunostimulant; vaccine; immunotherapy;  
 KW anti-HCV vaccine; AIDS; hepatitis C virus; HCV; polyprotein precursor.  
 XX  
 OS Hepatitis C virus.  
 XX  
 FH Key Location/Qualifiers  
 FT Peptide 2594..2602  
 FT /note= "specifically claimed peptide in claims 4 and 5"  
 XX  
 PN WO2004071414-A2.  
 XX



```
Db      181 LLSCLTVPAS 190

RESULT 14
AAR14557
ID AAR14557 standard; protein; 215 AA.
XX
AC AAR14557;
XX
XX 24-OCT-2003 (revised)
DT 24-JAN-1992 (first entry)
XX
DE HCV Capsid protein encoded by clone 56.
XX
XX non-A, non-B hepatitis; PT-NANBH; liver disease.
XX
XX Hepatitis C virus; Virus.
XX
XX WO9115516-A.
XX
PD 17-OCT-1991.
XX
XX 06-APR-1990; 90US-00505611.
XX
XX 06-APR-1990; 90US-00505611.
PR 09-OCT-1990; 90US-00594854.
XX
XX (GENE-) GENELABS INC.
XX
XX Reyes G, Kim JP, Moeckli R, Simonsen CC;
XX
XX WPI; 1991-325174/44.
DR N-PSDB; AAQ14279.
XX
XX Hepatitis C virus epitope(s) immuno-reactive with HCV infected sera
PT - useful for detection of HCV infections and as HCV vaccine.
XX
XX Claim 14; Page 110; 150pp; English.
XX
XX Putative HCV capsid protein clones were identified by immunoscreening.
CC Clone 56 was sequenced and found to include a run of adenine residues
CC around nucleotides 25 to 34; such sequences are similar to sequences
CC known to promote translation frameshifting. The mol. wt. of the protein
CC deduced from the open reading frame is 23.5kD. (Updated on 24-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 215 AA;

Query Match          99.1%; Score 1031; DB 2; Length 215;
Best Local Similarity 99.0%; Pred. No. 2.2e-90;
Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db 1 MGTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
QY 61 RRQPIKARRPEGRWTAQPGYPWPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Db 61 RRQPIKARRPEGRWTAQPGYPWPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
QY 121 KVDTLTTCGFADLMGYIPAVGAPLGGAAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVDTLTTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

Search completed: August 19, 2005, 12:39:27
Job time : 119.5 secs
```

---

```
Db      181 LLSCLTVPAS 190

RESULT 14
AAR14557
ID AAR14557 standard; protein; 215 AA.
XX
AC AAR14557;
XX
XX 24-OCT-2003 (revised)
DT 24-JAN-1992 (first entry)
XX
DE HCV Capsid protein encoded by clone 56.
XX
XX non-A, non-B hepatitis; PT-NANBH; liver disease.
XX
XX Hepatitis C virus; Virus.
XX
XX WO9115516-A.
XX
PD 17-OCT-1991.
XX
XX 06-APR-1990; 90US-00505611.
XX
XX 06-APR-1990; 90US-00505611.
PR 09-OCT-1990; 90US-00594854.
XX
XX (GENE-) GENELABS INC.
XX
XX Reyes G, Kim JP, Moeckli R, Simonsen CC;
XX
XX WPI; 1991-325174/44.
DR N-PSDB; AAQ14279.
XX
XX Hepatitis C virus epitope(s) immuno-reactive with HCV infected sera
PT - useful for detection of HCV infections and as HCV vaccine.
XX
XX Claim 14; Page 110; 150pp; English.
XX
XX Putative HCV capsid protein clones were identified by immunoscreening.
CC Clone 56 was sequenced and found to include a run of adenine residues
CC around nucleotides 25 to 34; such sequences are similar to sequences
CC known to promote translation frameshifting. The mol. wt. of the protein
CC deduced from the open reading frame is 23.5kD. (Updated on 24-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 215 AA;

Query Match          99.1%; Score 1031; DB 2; Length 215;
Best Local Similarity 99.0%; Pred. No. 2.2e-90;
Matches 189; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
Db 1 MGTNPKQKKNKNTNRRPDQVFPGGQIVGGVYLLPRGPRLGVRATRKTSERSQPRG 60
QY 61 RRQPIKARRPEGRWTAQPGYPWPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
Db 61 RRQPIKARRPEGRWTAQPGYPWPLYGNECGWAGWLLSPRGRSPSWGPTDPRRSRNLG 120
QY 121 KVDTLTTCGFADLMGYIPAVGAPLGGAAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
Db 121 KVDTLTTCGFADLMGYIPLVGAPLGGAAARALAHGVRLVDGVNYATGNLPGCSFSIFLLA 180
QY 181 LLSCLTVPASA 191
Db 181 LLSCLTVPASA 191

RESULT 15
AAR14558
ID AAR14558 standard; protein; 215 AA.
XX
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2005, 13:22:28 ; Search time 165 Seconds  
(without alignments)

27.932 Million cell updates/sec

Title: US-10-770-117-1

Perfect score: 47

Sequence: 1 DLGYPAP 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 68540

Minimum DB seq length: 0  
Maximum DB seq length: 50

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	59.6	50	Q9C1W4	Q9C1W4 schizosacch
2	27	57.4	27	Q66269	Q66269 cucumber mo
3	27	57.4	27	Q66270	Q66270 cucumber mo
4	27	57.4	27	Q89492	Q89492 cucumber mo
5	27	57.4	32	Q7KYP7	Q7KYP7 homo sapien
6	27	57.4	46	Q8WTQ2	Q8WTQ2 homo sapien
7	27	57.4	48	Q9P4B0	Q9P4B0 letharia vu
8	26	55.3	16	Q8R963	Q8R963 helicobacte
9	26	55.3	19	Q9UCH5	Q9UCH5 homo sapien
10	26	55.3	24	Q94370	Q94370 caenorhabdi
11	26	55.3	38	Q8EG71	Q8EG71 shewanella
12	26	55.3	40	Q9PA42	Q9PA42 xylella fas
13	26	55.3	41	Q663D0	Q663D0 yersinia ps
14	26	55.3	44	Q7QZ58	Q7QZ58 giardia lam
15	26	55.3	44	Q9ZG06	Q9ZG06 escherichia
16	26	55.3	46	Q9WP35	Q9WP35 tt virus. p
17	26	55.3	50	Q6AIU2	Q6AIU2 desulfotale
18	25	53.2	19	Q84859	Q84859 unidentifie
19	25	53.2	21	Q7M035	Q7M035 rattus norv
20	25	53.2	24	Q7M1L8	Q7M1L8 phleum prat
21	25	53.2	27	PSAF_SYNVU	P23078 synechococc
22	25	53.2	31	Q8Z612	Q8Z612 salmonella
23	25	53.2	32	Q8ES42	Q8ES42 rattus norv
24	25	53.2	36	Q86SS5	Q86SS5 homo sapien
25	25	53.2	39	Q866I9	Q866I9 callicebus
26	25	53.2	39	Q866J0	Q866J0 micopithecus
27	25	53.2	39	Q866J1	Q866J1 ateles geof
28	25	53.2	39	Q866J2	Q866J2 macaca mula
29	25	53.2	39	Q866J3	Q866J3 naealis lar
30	25	53.2	39	Q866J4	Q866J4 aotus trivi
31	25	53.2	39	Q866J5	Q866J5 callithrix

32	25	53.2	39	2	Q866J6	Q866J6 alouatta se
33	25	53.2	39	2	Q866J7	Q866J7 presbytis e
34	25	53.2	39	2	Q866J8	Q866J8 cercopithec
35	25	53.2	39	2	Q866J9	Q866J9 gorilla gor
36	25	53.2	39	2	Q866K0	Q866K0 hyllobates l
37	25	53.2	39	2	Q866K1	Q866K1 mandrillus
38	25	53.2	39	2	Q866K2	Q866K2 pan troglod
39	25	53.2	39	2	Q866K3	Q866K3 allenopithe
40	25	53.2	40	2	Q91K22	Q91K22 hepatitis c
41	25	53.2	43	2	Q72BP3	Q72BP3 desulfovibr
42	25	53.2	44	1	N160_XENLA	P83722 xenopus lae
43	25	53.2	47	2	Q7UYN7	Q7UYN7 rhodopirell
44	25	53.2	47	1	Y16H_BPT4	P39512 bacterioph
45	25	53.2	49	2	Q64CG7	Q64CG7 uncultured
46	24.5	52.1	47	2	Q9CW87	Q9CW87 mus musculu
47	24	51.1	18	2	Q6LD16	Q6LD16 homo sapien
48	24	51.1	25	2	Q9BG19	Q9BG19 cheirogaleu
49	24	51.1	25	2	Q9WMI4	Q9WMI4 pseudomonas
50	24	51.1	27	2	Q9TZ20	Q9TZ20 caenorhabdi
51	24	51.1	29	2	Q93DQ1	Q93DQ1 uncultured
52	24	51.1	33	2	Q6UAS5	Q6UAS5 bacterioph
53	24	51.1	36	1	NLTP_PINPI	P26912 pinus pinea
54	24	51.1	36	1	PSAL_CVACA	Q9TM24 cyanidium c
55	24	51.1	36	2	Q8CTU7	Q8CTU7 staphylococ
56	24	51.1	36	2	Q35355	Q35355 rattus norv
57	24	51.1	38	2	Q886X3	Q886X3 pseudomonas
58	24	51.1	40	2	Q879L6	Q879L6 streptococ
59	24	51.1	42	2	Q7MRD6	Q7MRD6 wolinnella s
60	24	51.1	45	2	Q8FY22	Q8FY22 escherichia
61	24	51.1	46	2	Q8EHZ6	Q8EHZ6 shewanella
62	24	51.1	47	2	Q86UL5	Q86UL5 homo sapien
63	24	51.1	49	2	Q6EEJ0	Q6EEJ0 sus scrofa
64	24	51.1	49	2	Q8S0U2	Q8S0U2 shigella so
65	23	48.9	9	2	Q88612	Q88612 middelburg
66	23	48.9	12	1	RF1_CONSP	P58805 conus spuri
67	23	48.9	16	2	Q7M263	Q7M263 escale cere
68	23	48.9	18	2	Q9S8Y9	Q9S8Y9 poa pratens
69	23	48.9	20	2	Q7M262	Q7M262 holcus lana
70	23	48.9	20	2	Q7M264	Q7M264 lolium pere
71	23	48.9	20	2	Q9S8X9	Q9S8X9 phleum prat
72	23	48.9	20	2	Q9S8Y0	Q9S8Y0 phleum prat
73	23	48.9	23	2	Q6JDN4	Q6JDN4 canis fami
74	23	48.9	25	2	Q8LTE8	Q8LTE8 bacterioph
75	23	48.9	25	2	Q7PCD2	Q7PCD2 bacterioph
76	23	48.9	30	2	Q9TQ11	Q9TQ11 equus cabal
77	23	48.9	31	2	Q9QYV4	Q9QYV4 mus musculu
78	23	48.9	33	2	Q15747	Q15747 homo sapien
79	23	48.9	33	2	Q9KSG4	Q9KSG4 vibrio chol
80	23	48.9	34	2	Q13871	Q13871 homo sapien
81	23	48.9	34	2	Q9RCC0	Q9RCC0 yersinia ps
82	23	48.9	34	2	Q9RCD1	Q9RCD1 yersinia pe
83	23	48.9	34	2	Q8F1R4	Q8F1R4 leptospira
84	23	48.9	34	2	Q9W7P6	Q9W7P6 morone saxa
85	23	48.9	35	2	Q7M3X7	Q7M3X7 geukensia d
86	23	48.9	35	2	Q18877	Q18877 canis fami
87	23	48.9	35	2	Q9TSX6	Q9TSX6 sus scrofa
88	23	48.9	35	2	Q7TTL5	Q7TTL5 prochloroco
89	23	48.9	35	2	Q7TTL7	Q7TTL7 prochloroco
90	23	48.9	35	2	Q9YNI2	Q9YNI2 human immun
91	23	48.9	36	2	Q7QV88	Q7QV88 giardia lam
92	23	48.9	36	2	Q9NGN1	Q9NGN1 strongyloce
93	23	48.9	37	2	Q8EDY3	Q8EDY3 shewanella
94	23	48.9	38	2	Q7M0Q2	Q7M0Q2 xanthomonas
95	23	48.9	38	2	Q9JUP1	Q9JUP1 neisseria m
96	23	48.9	43	2	Q722M6	Q722M6 listeria m
97	23	48.9	44	2	Q83YL2	Q83YL2 bacillus th
98	23	48.9	44	2	Q9REI7	Q9REI7 helicobacte
99	23	48.9	44	2	Q7V203	Q7V203 prochloroco
100	23	48.9	45	2	Q84WH1	Q84WH1 arabidopsis

ALIGNMENTS

```

RESULT 1
ID Q9C1W4 PRELIMINARY; PRT; 50 AA.
AC Q9C1W4;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE SPBC713.11c protein.
GN Name=SPBC713.11c;
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972h-;
RC MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Gobie A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S., Warren T., Whitehead S.,
RA Woodward J., Volckert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schafer M., Muller-Auer S.,
RA Gabel C., Fuchs M., Dusterhoft A., Fritz C., Holzer E., Moestl D.,
RA Hilbert H., Borzym K., Langer I., Beck A., Leirach H., Reinhardt R.,
RA Pohl T.M., Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Drano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
DR EMBL; AL512943; CAC22612.1; -.
DR GeneDB SPombe; SPBC713.11c; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR000612; UPF0057.
DR Pfam; PF01679; UPF0057; 1.
DR PROSITE; PS01309; UPF0057; 1.
SQ SEQUENCE 50 AA; 5475 MW; 503BABFECC027EF2 CRC64;

Query Match 59.6%; Score 28; DB 2; Length 50;
Best Local Similarity 42.9%; Pred. No. 8.6e+02;
Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 MGYPAP 9
Db 34 LGYVPGI 40

RESULT 2
Q66269
ID Q66269 PRELIMINARY; PRT; 27 AA.
AC Q66269;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Cucurbit mosaic virus open reading frame, complete cds, and satellite RNA.
DE DE
OS Cucurbit mosaic virus (cucurbit mosaic cucumovirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;

```

```

OC Cucumovirus.
OX NCBI_TaxID=12305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=F;
RC MEDLINE=90171930; PubMed=1689769;
RA Hidaka S., Hanada K., Ishikawa K.;
RT "In vitro messenger properties of a satellite RNA of cucumber mosaic virus."
RL J. Gen. Virol. 71:439-442(1990).
DR EMBL; D00699; BAA00605.1; -.
DR J. Gen. Virol. 71:439-442(1990).
SQ SEQUENCE 27 AA; 2874 MW; 1DDA41BF3FAL328A CRC64;

Query Match 57.4%; Score 27; DB 2; Length 27;
Best Local Similarity 66.7%; Pred. No. 7.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MGYPAP 8
Db 19 VGYLPA 24

RESULT 3
Q66270
ID Q66270 PRELIMINARY; PRT; 27 AA.
AC Q66270;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
DE Cucurbit mosaic virus (cucurbit mosaic cucumovirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Cucumovirus.
OX NCBI_TaxID=12305;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=E;
RC MEDLINE=88219519; PubMed=3369085;
RA Hidaka S., Hanada K., Ishikawa K., Miura K.;
RT "Complete nucleotide sequence of two new satellite RNAs associated with cucumber mosaic virus."
RL Virology 164:326-333(1988).
DR EMBL; M20844; AAA66611.1; -.
RW Hypothetical protein.
SQ SEQUENCE 27 AA; 2783 MW; 1D80426D890A462F CRC64;

Query Match 57.4%; Score 27; DB 2; Length 27;
Best Local Similarity 66.7%; Pred. No. 7.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MGYPAP 8
Db 19 VGYLPA 24

RESULT 4
Q89492
ID Q89492 PRELIMINARY; PRT; 27 AA.
AC Q89492;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ORF 1.
OS Cucurbit mosaic virus (cucurbit mosaic cucumovirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC Cucumovirus.
OX NCBI_TaxID=12305;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=88179532; PubMed=3354198;
RA Kaper J.M., Tounignant M.E., Steen M.T.;
RT "Cucurbit mosaic virus-associated RNA 5: XI. Comparison of 14 CARNA 5 variants relates ability to induce tomato necrosis to a conserved

```

RT nucleotide sequence."  
 RL Virology 163:284-292(1988).  
 DR EMBL; M20357; AAA46403.1; -  
 DR EMBL; M20350; AAA46385.1; -  
 DR EMBL; M20352; AAA46388.1; -  
 DR EMBL; M20353; AAA46391.1; -  
 DR EMBL; M20354; AAA46394.1; -  
 DR EMBL; M20355; AAA46397.1; -  
 DR EMBL; M20356; AAA46400.1; -  
 DR PIR; JE0015; JE0015.  
 DR PIR; JS0104; JS0104.  
 SQ SEQUENCE 27 AA; 2827 MW; BC60426D890A589A CRC64;  
 Query Match 57.4%; Score 27; DB 2; Length 27;  
 Best Local Similarity 66.7%; Pred. No. 7.9e+02;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 MGYIPA 8  
 DB 19 VGYLPA 24  
 RESULT 5  
 Q7KYP7 PRELIMINARY; PRT; 32 AA.  
 AC Q7KYP7;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Dystrophin C-terminal region (32 AA) (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=8181947; PubMed=2648158; DOI=10.1038/338509a0;  
 RA Feener C.A., Koenig M., Kunkel L.M.;  
 RT "Alternative splicing of human dystrophin mRNA generates isoforms at the carboxy terminus."  
 RL Nature 338:509-511(1989).  
 DR EMBL; X15149; CAA33246.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 32 AA; 3856 MW; CC3917998F22F47E CRC64;  
 Query Match 57.4%; Score 27; DB 2; Length 32;  
 Best Local Similarity 80.0%; Pred. No. 9.2e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 MGYIP 7  
 DB 16 MGYLP 20  
 RESULT 6  
 Q8WTQ2 PRELIMINARY; PRT; 46 AA.  
 AC Q8WTQ2;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Dystrophin (Fragment).  
 GN Name=DMD;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Beggs A.H., Koenig M., Kunkel L.M.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
 RA den Dunnen J.T.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF213442; AAL61586.1; -  
 DR EMBL; AF213442; AAL61587.1; -  
 FT NON\_TER 1  
 FT NON\_TER 46  
 SQ SEQUENCE 46 AA; 5376 MW; DE4AC42C16208108 CRC64;  
 Query Match 57.4%; Score 27; DB 2; Length 46;  
 Best Local Similarity 80.0%; Pred. No. 1.3e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 MGYIP 7  
 DB 30 MGYLP 34  
 RESULT 7  
 Q9P4B0 PRELIMINARY; PRT; 48 AA.  
 AC Q9P4B0;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 OS Letharia vulpina.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Lecanoromycetes;  
 OC Lecanorales; Lecanorineae; Parmeliaceae; Letharia.  
 OX NCBI\_TaxID=129387;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21546980; PubMed=11686674; DOI=10.1006/fgbi.2001.1291;  
 RA "Outcrossing and recombination in the lichenized fungus Letharia."  
 RL Fungal Genet. Biol. 34:83-92(2001).  
 DR EMBL; AF228463; AAF87290.1; -  
 KW Hypothetical protein.  
 FT NON\_TER 48  
 SQ SEQUENCE 48 AA; 5779 MW; 9AFE570D10EP9755 CRC64;  
 Query Match 57.4%; Score 27; DB 2; Length 48;  
 Best Local Similarity 83.3%; Pred. No. 1.3e+03;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LMGYIP 7  
 DB 27 LKGYP 32  
 RESULT 8  
 Q9R963 PRELIMINARY; PRT; 16 AA.  
 AC Q9R963;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
 DE PepC (Fragment).  
 GN Name=pepC;  
 OS Helicobacter pylori (Campylobacter pylori).  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Helicobacteraceae; Helicobacter.  
 OX NCBI\_TaxID=210;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=F31;  
 RX MEDLINE=98453456; PubMed=9780260;  
 RA Ito Y., Azuma T., Ito S., Suto H., Miyaji H., Yamazaki Y., Kohli Y.,  
 RA Kuriyama M.;  
 RT "Full-length sequence analysis of the vacA gene from cytotoxic and noncytotoxic Helicobacter pylori."  
 RL J. Infect. Dis. 178:1391-1398(1998).  
 DR EMBL; AF049623; AAD04263.1; -

```

FT  NON TER      1      1
FT  NON TER      16     16
SQ  SEQUENCE    16 AA; 1938 MW;  3C33D403ABE61428 CRC64;

Query Match      55.3%; Score 26; DB 2; Length 16;
Best Local Similarity 62.5%; Pred. No. 7.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy  2  LMGYIPAV 9
    | : ||| |
Db  1  LIAIPIQV 8

RESULT 9
ID  Q9UCH5      PRELIMINARY;      PRT;      19 AA.
AC  Q9UCH5;
DT  01-MAY-2000 (TrEMBLrel. 13, Created)
DT  01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Natural killer enhancing factor, NKEF (fragment).
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE.
RX  MEDLINE=93215001; PubMed=8462106; DOI=10.1006/cimm.1993.1043;
RA  Shau H., Gupta R.K., Golub S.H.;
RT  "Identification of a natural killer enhancing factor (NKEF) from human
RT  erythroid cells.";
RL  Cell. Immunol. 147:1-11(1993).
DR  GO; GO:0005829; C:cytosol; NAS.
FT  NON TER      1
FT  NON TER      19
SQ  SEQUENCE    19 AA; 2012 MW;  69DC04F5DD5BC56F CRC64;

Query Match      55.3%; Score 26; DB 2; Length 19;
Best Local Similarity 62.5%; Pred. No. 9.1e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy  1  DLMGYIPA 8
    | : ||| |
Db  9  DEMGEVPA 16

RESULT 10
ID  Q94370      PRELIMINARY;      PRT;      24 AA.
AC  Q94370;
DT  01-FEB-1997 (TrEMBLrel. 02, Created)
DT  01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  PPI-like Ser/Thr protein phosphatase (fragment).
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC  Rhabditidae; Peloderinae; Caenorhabditis.
OX  NCBI_TaxID=6239;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Zeke T., Gergely P., Dombradi V.;
RL  Submitted (JUL-1996) to the EMBL/GenBank/DBSJ databases.
DR  EMBL; 277730; CAB01289.1; -.
DR  PIR; T42390; T42390.
FT  NON TER      1
FT  NON TER      24
SQ  SEQUENCE    24 AA; 2787 MW;  004023AE033FCD01 CRC64;

Query Match      55.3%; Score 26; DB 2; Length 24;
Best Local Similarity 55.6%; Pred. No. 1.1e+03;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy  1  DLMGYIPAV 9
    | : ||| |

```

```

Db  10  DAVGYPPKV 18
    | : ||| |

RESULT 11
ID  Q8EG71      PRELIMINARY;      PRT;      38 AA.
AC  Q8EG71;
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT  01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE  Hypothetical protein SO1737.
GN  OrderedLocustNames=SO1737;
OS  Shewanella oneidensis.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC  Shewanellaceae; Shewanella.
OX  NCBI_TaxID=70863;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=MR-1;
RX  MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA  Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA  Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA  Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA  Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA  Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,
RA  White O., Wolf A.M., Vamathevan J.O., Weidman J.F., Imbraim M.,
RA  Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA  Uterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA  Venter J.C., Nealon K.H., Fraser C.M.;
RT  "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT  Shewanella oneidensis.";
RL  Nat. Biotechnol. 20:1118-1123(2002).
DR  EMBL; AE015618; AAN54791.1; -.
DR  TIGR; SO1737; -.
KW  Complete proteome; Hypothetical protein.
SQ  SEQUENCE    38 AA; 4405 MW;  410074CB8F607359 CRC64;

Query Match      55.3%; Score 26; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.7e-03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2  LMGYI 6
    | ||| |
Db  5  LMGYI 9

RESULT 12
ID  Q9PA42      PRELIMINARY;      PRT;      40 AA.
AC  Q9PA42;
DT  01-OCT-2000 (TrEMBLrel. 15, Created)
DT  01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT  01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE  Hypothetical protein.
GN  OrderedLocustNames=Xf2678;
OS  Xylella fastidiosa.
OC  Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC  Xanthomonadaceae; Xylella.
OX  NCBI_TaxID=2371;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=9a5c;
RX  MEDLINE=20365717; PubMed=10910347; DOI=10.1038/35018003;
RA  Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA  Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA  Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA  Bueno M.R.P., Camargo A.A., Canargo L.E.A., Carraro D.M., Carrer H.,
RA  Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA  Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA  Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA  Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA  Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,

```

```

RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Martins E.M.F., Marino C.L.,
RA Marques M.V., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Menck C.F.M., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Moon D.H., Nagai M.A., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Palmieri D.A., Paris A.,
RA de Oliveira M.C., de Oliveira R.C., Pereira H.A. Jr., Pesquero J.B.,
RA Peixoto B.R., Pereira G.A.G., Rodrigues V., de Rosa A.J.M.,
RA Quaggio R.B., Roberto P.G., Santelli R.V., Sawasaki H.E.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RA "The genome sequence of the plant pathogen Xylella fastidiosa."
RL Nature 406:151-159(2000).
DR EMBL: AE004073; AAF85475.1; -.
DR PIR: A82529; A82529.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 40 AA; 4601 MW; 4C42AF68E5999933 CRC64;

Query Match 55.3%; Score 26; DB 2; Length 40;
Best Local Similarity 50.0%; Pred. No. 1.8e+03;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGYIPAV 9
Db 3 LLGYVESV 10

RESULT 13
Q663D0 PRELIMINARY; PRT; 41 AA.
AC Q663D0;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=pYtcb0024;
OS Yersinia pseudotuberculosis IP 32953.
OG Plasmid pYtcb32953.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=273123;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP 32953;
RX PubMed=15358858;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin L.V.,
RA Brubaker R.R., Fowler J., Hinnebusch B.J., Marceau M., Medigue C.,
RA Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RA "Insights into the genome evolution of Yersinia pestis through whole
RT genome comparison with Yersinia pseudotuberculosis."
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL: BX936400; CAF25466.1; -.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 41 AA; 4704 MW; 89DC6E905AE923F0 CRC64;

Query Match 55.3%; Score 26; DB 2; Length 41;
Best Local Similarity 71.4%; Pred. No. 1.8e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LMGYIPAV 7
Db 15 LLGYIPAV 21

RESULT 14

```

```

Q70Z58
ID Q70Z58 PRELIMINARY; PRT; 44 AA.
AC Q70Z58;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GUP_22_47461_47327.
OS Giardia lamblia ATCC 50803.
OC Eukaryota; Diplomonadida; Hexamitidae; Giardiinae; Giardia.
OX NCBI_TaxID=184922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB CG;
RA Morrison H.G., McArthur A.G., Adam R.D., Aley S.B., Gillin F.D.,
RA Olsen G.J., Sogin M.L.;
RA "Draft sequence of the Giardia lamblia genome."
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL: AACB01000045; EAA40351.1; -.
SQ SEQUENCE 44 AA; 5425 MW; 6CCAB8846627693A CRC64;

Query Match 55.3%; Score 26; DB 2; Length 44;
Best Local Similarity 71.4%; Pred. No. 1.9e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 MGYPVAV 9
Db 1 MGLIPSV 7

RESULT 15
Q9ZGU6 PRELIMINARY; PRT; 44 AA.
ID Q9ZGU6;
AC Q9ZGU6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Regulatory protein.
GN OrderedlocusNames=L7024;
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
RX MEDLINE=98391744; PubMed=9722640; DOI=10.1093/nar/26.18.4196;
RA Burland V., Shao Y., Perna N.T., Plunkett G. III, Sofia H.J.,
RA Blattner F.R.;
RA "The complete DNA sequence and analysis of the large virulence plasmid
RT of Escherichia coli O157:H7."
RL Nucleic Acids Res. 26:4196-4204 (1998).
DR EMBL: AF074613; AAC70092.1; -.
DR PIR: T42124; T42124.
DR InterPro: IPR011010; DNA_brk_join_enz.
KW Complete proteome; Plasmid.
SQ SEQUENCE 44 AA; 4951 MW; 824EF81ED8273816 CRC64;

Query Match 55.3%; Score 26; DB 2; Length 44;
Best Local Similarity 62.5%; Pred. No. 1.9e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LMGYIPAV 9
Db 12 LMGYLGV 19

RESULT 16
Q9WP35 PRELIMINARY; PRT; 46 AA.
ID Q9WP35

```

```

AC Q9WP35;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Putative core/capsid protein (Fragment).
OS TT virus.
OC Viruses; ssDNA viruses; Anellovirus.
OX NCBI_TaxID=69887;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93350008; PubMed=10423145;
RA Ball J K., Curran R., Berridge S., Grabowska A.M., Jameson C.L.,
RA Thomson B.J., Irving W.L., Sharp P.M.;
RT "TT virus sequence heterogeneity in vivo: evidence for co-infection
RT with multiple genetic types.";
RL J. Gen. Virol. 80:1759-1768(1999).
DR EMBL; AF145015; AAD42373.1; -.
DR InterPro; IPR004219; TTvirus_Unk.
DR Pfam; PF02956; TT_ORF1.1.
FT NON_TER 1
FT NON_TER 46
SQ SEQUENCE 46 AA; 5256 MW; 335B2E92048DFEC3 CRC64;

Query Match 55.3%; Score 26; DB 2; Length 46;
Best Local Similarity 66.7%; Pred. No. 2e+03;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGYIP 7
Db 35 LRGYVP 40

RESULT 17
ID Q6AIU2 PRELIMINARY; PRT; 50 AA.
AC Q6AIU2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Hypothetical protein.
GN OrderedLocusNames=DP3009;
OS Desulfotalea psychrophila.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfobacterales;
OC Desulfobulbaceae; Desulfotalea.
OX NCBI_TaxID=84980;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=LSv54 / DSM 12343;
RX PubMed=15305914;
RA Rabus R., Ruepp A., Frickey T., Rattei T., Partmann B., Stark M.,
RA Bauer M., Zibat A., Lombardot T., Becker I., Amann J., Gellner K.,
RA Teeling H., Leuschner W.D., Gloeckner F.-O., Lupas A.N., Amann R.,
RA Klenk H.-P.;
RT "The genome of Desulfotalea psychrophila, a sulfate-reducing bacterium
RT from permanently cold Arctic sediments.";
RL Environ. Microbiol. 6:887-902(2004).
DR EMBL; CR522870; CAG37738.1; -.
KW Complete proteome.
SQ SEQUENCE 50 AA; 5611 MW; 0882A7347C205CD4 CRC64;

Query Match 55.3%; Score 26; DB 2; Length 50;
Best Local Similarity 62.5%; Pred. No. 2.2e+03;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DLMGYIPA 8
Db 39 DLPGYNPS 46

RESULT 18
Q84859
ID Q84859 PRELIMINARY; PRT; 19 AA.
AC Q84859;

```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Unidentified human poliovirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=40278;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87169734; PubMed=3031313;
RA Kuge S., Saito I., Nomoto A.;
RT "Primary structure of poliovirus defective-interfering particle
RT genomes and possible generation mechanisms of the particles.";
RL J. Mol. Biol. 192:473-487(1986).
DR EMBL; M30216; AAA66825.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 19
SQ SEQUENCE 19 AA; 2127 MW; F4D2596021AAEA1A CRC64;

Query Match 53.2%; Score 25; DB 2; Length 19;
Best Local Similarity 37.5%; Pred. No. 1.4e+03;
Matches 3; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPA 8
Db 12 DILGFVSA 19

RESULT 19
Q7M035 PRELIMINARY; PRT; 21 AA.
AC Q7M035;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Protein kinase.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE.
RX MEDLINE=933737942; PubMed=8365414;
RA Brunati A.M., James P., Donella-Deana A., Matoskova B., Robbins K.C.,
RA Pinna L.A.;
RT "Isolation and identification of two proto-oncogene products related
RT to c-fgr and fyn in a tyrosine-protein-kinase fraction of rat
RT spleen.";
RL Eur. J. Biochem. 216:323-327(1993).
DR PIR; S35676; S35676.
DR HSSP; P06241; 1G83.
DR InterPro; IPR001452; SH3.
DR ProDom; PD000066; SH3; 1.
SQ SEQUENCE 21 AA; 2335 MW; 40213D1508152FF8 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 21;
Best Local Similarity 80.0%; Pred. No. 1.6e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8
Db 15 GYIPS 19

RESULT 20
Q7M1L8 PRELIMINARY; PRT; 24 AA.
AC Q7M1L8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

```

DE Pollen allergen Phl p V (Fragment).  
OS Phleum pratense (Common timothy).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Poales; Poaceae; Poaceae;  
OC Avenae; Phleum.  
OC NCBI\_TaxID=15957;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=91322736; PubMed=1863892;  
RA Matthies F., Lowenstein H.;  
RT "Group V allergens in grass pollens. I. Purification and  
RT characterization of the group V allergen from Phleum pratense pollen,  
RT Phl p V.";  
RL Clin. Exp. Allergy 21:297-307(1991).  
DR PIR; A61505; A61505. 1  
FT NON TER 24 24  
FT NON TER 24 24  
SQ SEQUENCE 24 AA; 2198 MW; 6232BAF9CC85C17A CRC64;  
  
Query Match 53.2%; Score 25; DB 2; Length 24;  
Best Local Similarity 66.7%; Pred. No. 1.8e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 3 MGYIPA 8  
Db 2 LGYIPA 7  
  
RESULT 21  
PSAF SYNNU STANDARD; PRT; 27 AA.  
AC P23078;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 01-NOV-1991 (Rel. 20, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Photosystem I reaction centre subunit III (PSI-F) (Fragment).  
GN Name=psaf;  
OS Synecococcus vulcanus (Thermosynechococcus vulcanus).  
OC Bacteria; Cyanobacteria; Chroococcales; Thermosynechococcus.  
OC NCBI\_TaxID=32053;  
RN [1]  
RP SEQUENCE  
RX MEDLINE=89338747; PubMed=2503399; DOI=10.1016/0014-5793(89)80971-8;  
RA Koike H., Ikeuchi M., Hiyana T., Inoue Y.;  
RT "Identification of photosystem I components from the cyanobacterium,  
RT Synecococcus vulcanus by N-terminal sequencing.";  
RL FEBS Lett. 253:257-263(1989).  
CC -1- FUNCTION: Probably participates in efficiency of electron transfer  
CC from plastocyanin to P700 (or cytochrome c553 in algae and  
CC cyanobacteria). This plastocyanin-docking protein contributes to  
CC the specific association of plastocyanin to PSI.  
CC -1- SIMILARITY: Belongs to the psaf family.  
DR InterPro; IPR003666; PSI\_Psaf.  
DR Pfam; PF02507; PSI\_Psaf.1.  
KW Direct protein sequencing; Photosynthesis; Photosystem I.  
FT NON TER 27 27  
FT NON TER 27 27  
SQ SEQUENCE 27 AA; 2685 MW; 30FFACDC3F58C44 CRC64;  
  
Query Match 53.2%; Score 25; DB 1; Length 27;  
Best Local Similarity 50.0%; Pred. No. 2e+03;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 DLMGYIPA 8  
Db 1 DVAGLVPA 8  
  
RESULT 22  
Q8Z612 PRELIMINARY; PRT; 31 AA.  
AC Q8Z612;  
DT 01-MAR-2002 (TReMBLrel. 20, Created)  
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)  
DE Hypothetical protein STV2049.  
GN OrderedLocusNames=STV2049;  
OS Salmonella typhi.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Salmonella.  
OC NCBI\_TaxID=601;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=CT18;  
RX MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;  
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
RA Churcher C.M., Mungall K.L., Bentley S.D., Holden M.R.G., Sebaihia M.,  
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,  
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
RA Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,  
RA Whitehead S., Barrrell B.G.;  
RT "Complete genome sequence of a multiple drug resistant Salmonella  
RT enterica serovar Typhi CT18.";  
RL Nature 413:848-852(2001).  
DR EMBL; AL627272; CAD05592.1; -;  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 31 AA; 3168 MW; E3B537F92EE34199 CRC64;  
  
Query Match 53.2%; Score 25; DB 2; Length 31;  
Best Local Similarity 66.7%; Pred. No. 2.2e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 2 LMGYIP 7  
Db 7 IRGYIP 12  
  
RESULT 23  
Q9ES42 PRELIMINARY; PRT; 32 AA.  
AC Q9ES42;  
DT 01-MAR-2001 (TReMBLrel. 16, Created)  
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Thyroid hormone receptor beta 3 (Fragment).  
GN Name=TRbeta;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OC NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=Sprague-Dawley;  
RX MEDLINE=20501244; PubMed=11046130;  
RX DOI=10.1128/MCB.20.22.8329-8342.2000;  
RA Williams G.R.;  
RT "Cloning and characterization of two novel thyroid hormone receptor  
RT beta isoforms.";  
RL Mol. Cell. Biol. 20:8329-8342(2000).  
DR EMBL; AF239915; AAC33351.1; -;  
DR GO; GO:0004872; F:receptor activity; IEA.  
KW Receptor.  
FT NON TER 32 32  
FT NON TER 32 32  
SQ SEQUENCE 32 AA; 3696 MW; BA134847AC9B5F42 CRC64;  
  
Query Match 53.2%; Score 25; DB 2; Length 32;  
Best Local Similarity 80.0%; Pred. No. 2.3e+03;  
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 4 GYIPA 8  
Db 24 GYIPS 28  
  
RESULT 24

```

Q86SS5
ID Q86SS5 PRELIMINARY; PRT; 36 AA.
AC Q86SS5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein MGC35402.
GN Name=MGC35402;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences."
RL Proc. Natl.. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RP [2];
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC046449; AA046449.1; -.
KW Hypothetical protein.
SQ SEQUENCE 36 AA; 4322 MW; 224A71F7D357A6D8 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 36;
Best Local Similarity 57.1%; Pred. No. 2.6e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 MGYIPAV 9
Db 1 MGIIPSI 7

RESULT 25
Q86619 PRELIMINARY; PRT; 39 AA.
ID Q86619;
AC Q86619;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Plasmidogen (Fragment).
GN Name=PLG;
OS Callicebus moloch (Dusky titi).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Callicebinae;
OC Callicebus.
OX NCBI_TaxID=9523;
RN [1];
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192768; AA019635.1; -.

```

```

DR HSSP; P00747; 1K10.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
SQ SEQUENCE 39 AA; 4300 MW; A8DD66BB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8
Db 35 GYIPS 39

RESULT 26
Q866J0 PRELIMINARY; PRT; 39 AA.
ID Q866J0;
AC Q866J0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Plasmidogen (Fragment).
GN Name=PLG;
OS Miopithecus talapoin (Talapoin) (Cercopithecus talapoin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Miopithecus.
OX NCBI_TaxID=36231;
RN [1];
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192767; AA019634.1; -.
DR HSSP; P00747; 1K10.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
SQ SEQUENCE 39 AA; 4286 MW; A8DBBDBB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8
Db 35 GYIPS 39

RESULT 27
Q866J1 PRELIMINARY; PRT; 39 AA.
ID Q866J1;
AC Q866J1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Plasmidogen (Fragment).
GN Name=PLG;
OS Ateles Geoffroyi (Black-handed spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateles.
OX NCBI_TaxID=9509;
RN [1];
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.

```



```

DR EMBL; AY192766; AAO19633.1; -.
DR HSSP; P00747; IKIO.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PRINTS; PR00018; KRINGLE.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4328 MW; 595466BB44393F10 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GYIPA 8
DB 35 GYIPS 39

RESULT 28
Q866J2 ID Q866J2 PRELIMINARY; PRT; 39 AA.
AC Q866J2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Plasminogen (Fragment).
GN Name=PLG;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192765; AAO19632.1; -.
DR HSSP; P00747; IKIO.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4286 MW; A8DBDBB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GYIPA 8
DB 35 GYIPS 39

RESULT 29
Q866J3 ID Q866J3 PRELIMINARY; PRT; 39 AA.
AC Q866J3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Plasminogen (Fragment).
GN Name=PLG;
OS Nasalis larvatus (Proboscis monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC Nasalis.
OX NCBI_TaxID=43780;
RN [1]
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192765; AAO19632.1; -.
DR HSSP; P00747; IKIO.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4286 MW; A8DBDBB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GYIPA 8
DB 35 GYIPS 39

RESULT 30
Q866J4 ID Q866J4 PRELIMINARY; PRT; 39 AA.
AC Q866J4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen (Fragment).
GN Name=PLG;
OS Aotus trivirgatus (Night monkey) (Douroucoul).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=9505;
RN [1]
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192763; AAO19630.1; -.
DR HSSP; P00747; IKIO.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4242 MW; A8DD66BB47090C02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GYIPA 8
DB 35 GYIPS 39

RESULT 31
Q866J5 ID Q866J5 PRELIMINARY; PRT; 39 AA.
AC Q866J5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen (Fragment).
GN Name=PLG;
OS Callithrix sp. DB-2003.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Callitrich.
OX NCBI_TaxID=217956;
RN [1]
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192764; AAO19631.1; -.
DR HSSP; P00747; IKIO.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PROSITE; PS00070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4286 MW; A8DBDBB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GYIPA 8
DB 35 GYIPS 39

```

RA Boffelli D., Lewis K.D., Rubin E.M.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY192762; AA019629.1; -.  
 DR HSSP; P00747; IKIO.  
 DR InterPro; IPR000001; Kringle.  
 DR Pfam; PF00051; Kringle; 1.  
 DR PRINTS; PR00018; KRINGLE.  
 DR SMART; SM00130; KR; 1.  
 DR PROSITE; PS50070; KRINGLE\_2; 1.  
 KW Kringle.  
 FT NON\_TER 1 1  
 FT NON\_TER 39 39  
 SQ SEQUENCE 39 AA; 4300 MW; A8DD66BB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;  
 Best Local Similarity 80.0%; Pred. No. 2.7e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8  
 ||||:  
 Db 35 GYIPS 39

RESULT 32  
 Q866J6 PRELIMINARY; PRT; 39 AA.  
 ID Q866J6;  
 AC Q866J6;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Plasmidogen (Fragment).  
 GN Name=PLG;  
 OS Alouatta seniculus (Red howler monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Alouattinae;  
 OC Alouatta.  
 OX NCBI\_TaxID=9503;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Boffelli D., Lewis K.D., Rubin E.M.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY192761; AA019628.1; -.  
 DR HSSP; P00747; IKIO.  
 DR InterPro; IPR000001; Kringle.  
 DR Pfam; PF00051; Kringle; 1.  
 DR PRINTS; PR00018; KRINGLE.  
 DR SMART; SM00130; KR; 1.  
 DR PROSITE; PS50070; KRINGLE\_2; 1.  
 KW Kringle.  
 FT NON\_TER 1 1  
 FT NON\_TER 39 39  
 SQ SEQUENCE 39 AA; 4328 MW; 595466BB44393F10 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;  
 Best Local Similarity 80.0%; Pred. No. 2.7e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8  
 ||||:  
 Db 35 GYIPS 39

RESULT 33  
 Q866J7 PRELIMINARY; PRT; 39 AA.  
 ID Q866J7;  
 AC Q866J7;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Plasmidogen (Fragment).  
 GN Name=PLG;  
 OS Presbytis entellus (Hanuman langur).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae; Colobinae;  
 OC Presbytis.  
 OX NCBI\_TaxID=9574;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Boffelli D., Lewis K.D., Rubin E.M.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY192760; AA019627.1; -.  
 DR HSSP; P00747; IKIO.  
 DR InterPro; IPR000001; Kringle.  
 DR Pfam; PF00051; Kringle; 1.  
 DR PROSITE; PS50070; KRINGLE\_2; 1.  
 KW Kringle.  
 FT NON\_TER 1 1  
 FT NON\_TER 39 39  
 SQ SEQUENCE 39 AA; 4286 MW; A8DBBDBB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;  
 Best Local Similarity 80.0%; Pred. No. 2.7e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8  
 ||||:  
 Db 35 GYIPS 39

RESULT 34  
 Q866J8 PRELIMINARY; PRT; 39 AA.  
 ID Q866J8;  
 AC Q866J8;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Plasmidogen (Fragment).  
 GN Name=PLG;  
 OS Cercopithecus aethiops (Green monkey) (Grivet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=9534;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Boffelli D., Lewis K.D., Rubin E.M.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY192759; AA019626.1; -.  
 DR HSSP; P00747; IKIO.  
 DR InterPro; IPR000001; Kringle.  
 DR Pfam; PF00051; Kringle; 1.  
 DR PROSITE; PS50070; KRINGLE\_2; 1.  
 KW Kringle.  
 FT NON\_TER 1 1  
 FT NON\_TER 39 39  
 SQ SEQUENCE 39 AA; 4286 MW; A8DBBDBB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;  
 Best Local Similarity 80.0%; Pred. No. 2.7e+03;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8  
 ||||:  
 Db 35 GYIPS 39

RESULT 35  
 Q866J9 PRELIMINARY; PRT; 39 AA.  
 ID Q866J9;  
 AC Q866J9;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Plasmidogen (Fragment).  
 GN Name=PLG;  
 OS Gorilla gorilla (gorilla).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192758; AA019625.1; -.
DR HSSP; P00747; IK10.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4286 MW; A8DBDBB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8
Db 35 GYIPS 39

RESULT 36
Q866K0 ID Q866K0 PRELIMINARY; PRT; 39 AA.
AC Q866K0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Plasminogen (Fragment).
GN Name=PLG;
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192757; AA019624.1; -.
DR HSSP; P00747; IK10.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4286 MW; A8DBDBB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8
Db 35 GYIPS 39

RESULT 37
Q866K1 ID Q866K1 PRELIMINARY; PRT; 39 AA.
AC Q866K1;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen (Fragment).
GN Name=PLG;
OS Mandrillus leucophaeus (Drill) (Papio leucophaeus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Mandrillus.
OX NCBI_TaxID=9568;
RN [1]
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192756; AA019623.1; -.
DR HSSP; P00747; IK10.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR SMART; SM00130; KR; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4334 MW; A8DBDBB58E8F202 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8
Db 35 GYIPS 39

RESULT 38
Q866K2 ID Q866K2 PRELIMINARY; PRT; 39 AA.
AC Q866K2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Plasminogen (Fragment).
GN Name=PLG;
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192755; AA019622.1; -.
DR HSSP; P00747; IK10.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle; 1.
DR PROSITE; PS50070; KRINGLE_2; 1.
KW Kringle.
FT NON_TER 1 1
FT NON_TER 39 39
SQ SEQUENCE 39 AA; 4286 MW; A8DBDBB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8
Db 35 GYIPS 39

RESULT 39
Q866K3 ID Q866K3 PRELIMINARY; PRT; 39 AA.
AC Q866K3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Plasminogen (Fragment).
GN Name=PLG;
OS Allenopithecus nigroviridis (Allen's swamp monkey).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecidae;
OX NCBI_TaxID=54135;
RN [1]
RP SEQUENCE FROM N.A.
RA Boffelli D., Lewis K.D., Rubin E.M.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY192754; A0019621.1; -.
DR HSSP; P00747; IK10.
DR InterPro; IPR000001; Kringle.
DR Pfam; PF00051; Kringle_1.
DR SMART; SM00130; KR_1.
DR PROSITE; PSS0070; KRINGLE_2; 1.
KW Kringle.
FT NON TER 1 1
FT NON TER 39 39
SQ SEQUENCE 39 AA; 4286 MW; A8DBDBB44393F02 CRC64;

Query Match 53.2%; Score 25; DB 2; Length 39;
Best Local Similarity 80.0%; Pred. No. 2.7e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8
Db 35 GYIPS 39

RESULT 40
Q91K22
ID Q91K22 PRELIMINARY; PRT; 40 AA.
AC Q91K22;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Polyprotein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2132446; PubMed=11431747; DOI=10.1053/jhep.2001.25512;
RA Gimenez-Barcons M., Franco S., Suarez X., Fornis X., Ampurdanes S.,
RA Puig-Basagot F., Sanchez-Fuero A., Barrera J.M., Llovet J.M.,
RA Bruix J., Sanchez-Tapias J.M., Rodes J., Saiz J.C.;
RT "High amino acid variability within the NS5A of hepatitis C virus
(RCV) is associated with hepatocellular carcinoma in patients with
HCV-related cirrhosis."
RT Hepatology 34:158-167(2001).
DR EMBL; AF379250; AAK63467.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR Pfam; PF01560; HCV NS1; 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON TER 1 1
FT NON TER 40 40
SQ SEQUENCE 40 AA; 4162 MW; F8ECD3FE433B812E CRC64;

Query Match 53.2%; Score 25; DB 2; Length 40;
Best Local Similarity 57.1%; Pred. No. 2.8e+03;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGYIPA 8
Db 4 LMGVPA 10

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2005, 13:24:58 ; Search time 39 Seconds  
(without alignments)

22.204 Million cell updates/sec

Title: US-10-770-117-1

Perfect score: 47

Sequence: 1 DLMGVIPAV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 11837

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : PIR\_79:\*

1: Pir1.\*

2: Pir2.\*

3: Pir3.\*

4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	27	57.4	27	JE0015	probable 2.8K prot
2	27	57.4	27	JS0104	probable 2.8K prot
3	26	55.3	24	T42390	phosphoprotein pho
4	26	55.3	40	A82529	hypothetical prote
5	26	55.3	44	T42124	probable regulator
6	25	53.2	21	S35676	protein kinase - r
7	25	53.2	24	A61505	pollen allergen ph
8	25	53.2	27	S05219	photosystem I chai
9	25	53.2	31	AF0736	hypothetical prote
10	24	51.1	16	B36300	T-cell receptor de
11	24	51.1	33	B36204	GTP-binding regula
12	24	51.1	36	S22222	probable lipid tra
13	24	51.1	49	T44523	hypothetical prote
14	23	48.9	16	S38292	30K allergen - rye
15	23	48.9	19	S71871	glutathione transf
16	23	48.9	20	S38288	50K allergen - per
17	23	48.9	20	S38294	32K allergen - com
18	23	48.9	20	S38291	30K allergen - vel
19	23	48.9	33	H82216	hypothetical prote
20	23	48.9	34	A55758	myosin - human (fr
21	23	48.9	35	A61077	glue protein - rib
22	23	48.9	38	S03339	immunoglobulin-bin
23	23	48.9	38	E81889	hypothetical prote
24	23	48.9	45	G95916	hypothetical exope
25	23	48.9	46	T40735	ltuA protein - chl
26	23	48.9	46	D81678	conserved hypothet
27	22	46.8	17	PH1312	Ig heavy chain DJ
28	22	46.8	20	A05332	serum amyloid P-co
29	22	46.8	20	S06149	photosystem I chai

30	22	46.8	21	2	S21772	lipoxigenase (EC 1
31	22	46.8	26	2	E27011	MHC class II-assoc
32	22	46.8	30	2	C43591	51K outer membrane
33	22	46.8	33	2	E84341	hypothetical prote
34	22	46.8	36	2	I46593	myosin - pig (frag
35	22	46.8	37	2	S03941	hydrogen dehydroge
36	22	46.8	42	2	A30010	cytochrome-c oxida
37	22	46.8	43	2	PS0359	hypothetical prote
38	22	46.8	44	2	C35156	dihydrolipoamide d
39	22	46.8	45	2	B35156	dihydrolipoamide d
40	22	46.8	48	2	S16924	cytochrome cc3 - D
41	22	46.8	48	2	S11639	hypothetical prote
42	22	46.8	49	2	I38067	nitric-oxide synth
43	21	44.7	24	2	S34405	adenylate kinase (
44	21	44.7	25	2	S03275	photosystem II 22K
45	21	44.7	29	2	G90719	hypothetical prote
46	21	44.7	29	2	T48910	KDPF protein (vali
47	21	44.7	29	2	T36654	probable small mem
48	21	44.7	34	2	F86055	hypothetical prote
49	21	44.7	35	2	S42673	adhesive protein -
50	21	44.7	37	2	A10933	coagulation factor
51	21	44.7	37	2	F82102	hypothetical prote
52	21	44.7	38	2	S32314	sinapine synthase
53	21	44.7	39	2	F97313	hypothetical prote
54	21	44.7	40	2	C69400	hypothetical prote
55	21	44.7	41	2	PK0046	GTP-binding protei
56	21	44.7	42	2	G90273	hypothetical prote
57	21	44.7	44	2	E82479	hypothetical prote
58	21	44.7	45	2	H70240	hypothetical prote
59	21	44.7	45	2	H97508	hypothetical prote
60	21	44.7	47	2	T37087	probable IS elemen
61	21	44.7	47	2	AF1165	Hypothetical CDS [
62	21	44.7	48	2	A38882	GTP-binding protei
63	21	44.7	48	2	S26102	outer membrane pro
64	21	44.7	48	2	A35156	dihydrolipoamide d
65	21	44.7	49	2	T06923	high light-inducib
66	21	44.7	49	2	T40034	probable mitochond
67	21	44.7	49	2	C90080	hypothetical prote
68	20.5	43.6	40	2	A43853	fructose-bisphosph
69	20	42.6	9	2	S13636	coat protein beta
70	20	42.6	15	2	A36279	chemoattractant pr
71	20	42.6	18	2	A35704	cytochrome P450 ol
72	20	42.6	19	2	A41668	integrin alpha-7 c
73	20	42.6	20	2	A61525	alpha-lactalbumin
74	20	42.6	21	2	I50535	calmodulin - elect
75	20	42.6	22	2	I50533	calmodulin - elect
76	20	42.6	22	2	S55308	glutathione transf
77	20	42.6	23	2	S23518	stromelysin (EC 3.
78	20	42.6	23	2	PC4030	rRNA endonuclease
79	20	42.6	25	2	A61457	alpha-glucosidase
80	20	42.6	30	2	G81031	hypothetical prote
81	20	42.6	30	2	S57227	proboscipedia prot
82	20	42.6	31	2	B98020	N-acetylneuraminat
83	20	42.6	31	2	H98061	N-acetylneuraminat
84	20	42.6	33	2	PS0433	dyetrophin - rabbi
85	20	42.6	34	2	S34265	hypothetical prote
86	20	42.6	36	1	PIFNIT	histidyl-tRNA synt
87	20	42.6	36	2	T38913	photosystem I chai
88	20	42.6	36	2	S73308	photosystem I prot
89	20	42.6	36	2	S01531	photosystem I prot
90	20	42.6	37	1	S26087	plastoquinol-plast
91	20	42.6	37	2	S73239	conserved hypothet
92	20	42.6	37	2	D69283	Ivi protein III -
93	20	42.6	38	2	S70800	hypothetical prote
94	20	42.6	38	2	B71349	rex protein - simi
95	20	42.6	39	2	S44288	CAB/ElrIP/HLIP-rela
96	20	42.6	40	2	AH1862	hypothetical prote
97	20	42.6	40	2	E64039	hypothetical prote
98	20	42.6	41	2	T45906	hypothetical prote
99	20	42.6	42	2	T07474	omega-agatoxin III
100	20	42.6	43	2	E54252	

## ALIGNMENTS

## RESULT 1

JE0015  
 C:Species: cucurbit mosaic virus, CMV  
 C:Date: 31-Dec-1991 #sequence\_revision 28-May-1998 #text\_change 09-Jul-2004  
 C:Accession: JE0015; PS0286  
 R:Richards, K.E.; Jonard, G.; Jacquemond, M.; Lot, H.  
 Virolgy 89, 395-408, 1978  
 A:Title: Nucleotide sequence of cucumber mosaic virus-associated RNA 5.  
 A:Reference number: JE0015  
 A:Accession: JE0015  
 A:Status: translation not shown; conceptual translation of pseudogene  
 A:Molecule type: genomic RNA  
 A:Residues: 1-27 <RIC>  
 A:Cross-references: UNIPROT:Q89492  
 R:Kaper, J.M.; Tournant, M.E.; Steen, M.T.  
 Virology 163, 284-292, 1988  
 A:Title: Cucurbit mosaic virus-associated RNA 5. XI. Comparison of 14 CARNA 5 variants  
 A:Reference number: JE0018; MUID:88179532; PMID:3354198  
 A:Accession: PS0286  
 A:Status: translation not shown; conceptual translation of pseudogene  
 A:Molecule type: genomic RNA  
 A:Residues: 1-27 <MAP>  
 A:Cross-references: GB:M20350; NID:G331658; PIDN:AAA46385.1; PID:G331659  
 A:Note: this sequence is found in strain D and isolates Sq10, Ch20, X2NT3, X7, X12 and X  
 C:Genetics:  
 A:Map position: segment 5  
 C:Keywords: pseudogene

Query Match 57.4%; Score 27; DB 4; Length 27;  
 Best Local Similarity 66.7%; Pred. No. 94;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MGYIPA 8

Db 19 VGYLPA 24

## RESULT 2

JS0104  
 Probable 2.8K protein pseudogene - cucumber mosaic virus (strain Y)  
 C:Species: cucumber mosaic virus, CMV  
 A:Note: host Nicotiana tabacum cv. Xanthi nc (tobacco)  
 C:Date: 31-Mar-1992 #sequence\_revision 28-May-1998 #text\_change 09-Jul-2004  
 C:Accession: JS0104  
 R:Hidaka, S.; Ishikawa, K.; Takanami, Y.; Kubo, S.; Miura, K.  
 FEBS Lett. 174, 38-42, 1984  
 A:Title: Complete nucleotide sequence of RNA 5 from cucumber mosaic virus (strain Y).  
 A:Reference number: JA0077  
 A:Accession: JS0104  
 A:Status: translation not shown; conceptual translation of pseudogene  
 A:Molecule type: genomic RNA  
 A:Residues: 1-27 <HID>  
 A:Cross-references: UNIPROT:Q89492  
 C:Comment: CMV is an icosahedral virus containing a tripartite, positive-stranded RNA ge  
 nent referred to as satellite RNA.  
 C:Genetics:  
 A:Map position: segment 5  
 C:Keywords: pseudogene

Query Match 57.4%; Score 27; DB 4; Length 27;  
 Best Local Similarity 66.7%; Pred. No. 94;  
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MGYIPA 8

Db 19 VGYLPA 24

## RESULT 3

T42124  
 Probable regulatory protein - Escherichia coli plasmid p0157  
 C:Species: Escherichia coli

## T42390

phosphoprotein phosphatase (EC 3.1.3.16) - Caenorhabditis elegans (fragment)  
 C:Species: Caenorhabditis elegans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C:Accession: T42390  
 R:Zeke, T.; Gergely, P.; Dombradi, V.  
 submitted to the EMBL Data Library, July 1996  
 A:Description: The catalytic subunits of Ser/Thr protein phosphatases from Caenorhabditis  
 A:Reference number: Z22131  
 A:Accession: T42390  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: DNA  
 A:Residues: 1-24 <ZEK>  
 A:Cross-references: UNIPROT:Q94370; EMBL:Z77730; PIDN:CAE01289.1  
 C:Superfamily: serine/threonine protein phosphatase; phosphoesterase core homology; phosph  
 C:Keywords: phosphoric monoester hydrolase

Query Match 55.3%; Score 26; DB 2; Length 24;  
 Best Local Similarity 55.6%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9

Db 10 DAVGYPPKV 18

## RESULT 4

AB2529  
 hypothetical protein XF2678 [imported] - Xylella fastidiosa (strain 9a5c)  
 C:Species: Xylella fastidiosa  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C:Accession: AB2529  
 R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
 Nature 406, 151-157, 2000  
 A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A:Reference number: AB2515; MUID:20365717; PMID:10910347  
 A:Note: for a complete list of authors see reference number A59328 below  
 A:Accession: AB2529  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-40 <SIM>  
 A:Cross-references: UNIPROT:Q9PA42; GB:AE004073; GB:AE003849; NID:g9107904; PIDN:AAF8547;  
 A:Experimental source: strain 9a5c  
 R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Aj  
 Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.  
 as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohne  
 J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuranae, E.E.; Laig  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E.  
 A:Authors: Martins, E.M.F.; Matsumura, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki  
 A:Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira  
 M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
 A:Reference number: A59328  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: XF2678

Query Match 55.3%; Score 26; DB 2; Length 40;  
 Best Local Similarity 50.0%; Pred. No. 2.2e+02;  
 Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGYIPAV 9

Db 3 LLGYVESV 10

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C/Accession: T42124  
 R/Burland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.  
 Nucleic Acids Res. 26, 4196-4204, 1998  
 A/Title: The complete DNA sequence and analysis of the large virulence plasmid of *Escherichia coli* O157:H7  
 A/Reference number: 222068; MUID:98391744; PMID:9722640  
 A/Accession: T42124  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-44 <BR>  
 A/Cross-references: UNIPROT:Q9ZGU6; EMBL:AF074613; PIDN:AAC70092.1  
 A/Experimental source: strain BDL933; serotype O157:H7  
 C/Genetics:  
 A/Genome: plasmid pO157  
 A/Note: L7024  
 C/Superfamily: type 1 fimbriae regulatory protein fimE  
 C/Keywords: transcription regulation

Query Match 55.3%; Score 26; DB 2; Length 44;  
 Best Local Similarity 62.5%; Pred. No. 2.4e+02;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LMGYIPAV 9  
 |||||  
 Db 12 LMGYLGV 19

RESULT 6  
 S35676  
 protein kinase - rat  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 10-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
 C/Accession: S35676  
 R/Brunati, A.M.; James, P.; Donella-Deana, A.; Matoskova, B.; Robbins, K.C.; Pinna, L.A.  
 Eur. J. Biochem. 216, 323-327, 1993  
 A/Title: Isolation and identification of two proto-oncogene products related to c-fgr an  
 A/Reference number: S35675; MUID:93373942; PMID:8365414  
 A/Accession: S35676  
 A/Molecule type: protein  
 A/Residues: 1-21 <BR>  
 A/Cross-references: UNIPROT:Q7M035

Query Match 53.2%; Score 25; DB 2; Length 21;  
 Best Local Similarity 80.0%; Pred. No. 1.8e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8  
 |||||  
 Db 15 GYIPS 19

RESULT 7  
 A61505  
 pollen allergen Phl p v - common timothy (fragment)  
 C/Species: Phleum pratense (common timothy)  
 C/Date: 15-Oct-1994 #sequence\_revision 15-Oct-1994 #text\_change 09-Jul-2004  
 C/Accession: A61505  
 R/Matthiesen, F.; Lowenstein, H.  
 Clin. Exp. Allergy 21, 297-307, 1991  
 A/Title: Group V allergens in grass pollens. I. Purification and characterization of the  
 A/Reference number: A61505; MUID:91322736; PMID:1863892  
 A/Accession: A61505  
 A/Status: preliminary  
 A/Molecule type: protein  
 A/Residues: 1-24 <MAT>  
 A/Cross-references: UNIPROT:Q7MIL8  
 A/Note: 2-Ala, 5-Gly, and 15-Glu were also found  
 C/Superfamily: grass pollen allergen IX  
 C/Keywords: hydroxyproline; pollen  
 P:6,9,12,18,21,24/Modified site: hydroxyproline (Pro) #status experimental

Query Match 53.2%; Score 25; DB 2; Length 24;  
 Best Local Similarity 66.7%; Pred. No. 2.1e+02;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 MGYIPA 8  
 :|||  
 Db 2 LGIAPA 7

RESULT 8  
 S05219  
 photosystem I chain III - *Synechococcus* sp. (fragment)  
 N/Alternate names: photosystem I L4K protein; photosystem I plastocyanin-binding chain  
 C/Species: *Synechococcus* sp.  
 C/Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 30-Apr-1999  
 C/Accession: S05219  
 R/Koike, H.; Ikeuchi, M.; Hiayama, T.; Inoue, Y.  
 FEBS Lett. 253, 257-263, 1989  
 A/Title: Identification of photosystem I components from the cyanobacterium, *Synechococcus*  
 A/Reference number: S05218; MUID:89338747; PMID:2503399  
 A/Accession: S05219  
 A/Molecule type: protein  
 A/Residues: 1-27 <KOI>  
 A/Note: the authors identified this protein as photosystem I chain IV  
 R/Koike, H.; Ikeuchi, M.; Hiayama, T.; Inoue, Y.  
 FEBS Lett. 258, 180B, 1989  
 A/Reference number: S06842  
 A/Contents: annotation; erratum  
 A/Note: this is a revision to the identification of the sequence in reference S05218  
 C/Genetics:  
 A/Gene: psaf  
 C/Superfamily: photosystem I chain III  
 C/Keywords: photosynthesis; photosystem I; thylakoid

Query Match 53.2%; Score 25; DB 2; Length 27;  
 Best Local Similarity 50.0%; Pred. No. 2.4e+02;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DLMGYIPA 8  
 :|||  
 Db 1 DVAGLVEA 8

RESULT 9  
 AF0736  
 hypothetical protein STY2049 [imported] - *Salmonella enterica* subsp. *enterica* serovar Typh  
 C/Species: *Salmonella enterica* subsp. *enterica* serovar Typhi  
 A/Note: this species has also been called *Salmonella typhi*  
 C/Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002  
 C/Accession: AF0736  
 R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 A/Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* sero  
 A/Reference number: AB0502; MUID:21534947; PMID:11677608  
 A/Accession: AF0736  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-31 <PAR>  
 A/Cross-references: GB:AL513382; PIDN:CAD05592.1; PID:gl6503089; GSPDB:GN00176  
 C/Genetics:  
 A/Gene: STY2049

Query Match 53.2%; Score 25; DB 2; Length 31;  
 Best Local Similarity 66.7%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGYIP 7  
 :|||  
 Db 7 IRGYIP 12

RESULT 10

B36300  
T-cell receptor delta chain V-J region - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 30-May-1997  
C:Accession: B36300  
R:Matlis, L.A.; Fry, A.M.; Cron, R.Q.; Cotterman, M.M.; Dick, R.F.; Bluestone, J.A.  
Science 245, 746-749, 1989  
A>Title: Structure and specificity of a class II MHC alloreactive gammadelta T cell receptor  
A:Reference number: A36300; MUID:89368895; PMID:2528206  
C:Accession: B36300  
A>Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-16 <MAT>  
C:Keywords: T-cell receptor

Query Match 51.1%; Score 24; DB 2; Length 16;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIP 7  
Db 3 GYIP 6

RESULT 11  
D36204  
GTP-binding regulatory protein gamma-2 chain - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C>Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 19-Jul-1996  
C:Accession: D36204  
R:Gautam, N.; Northup, J.; Tamir, H.; Simon, M.I.  
Proc. Natl. Acad. Sci. U.S.A. 87, 7973-7977, 1990  
A>Title: G protein diversity is increased by associations with a variety of gamma subunits  
A:Reference number: A36204; MUID:91045919; PMID:2122451  
C:Accession: D36204  
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-33 <GAU>  
C:Superfamily: GTP-binding regulatory protein gamma chain

Query Match 51.1%; Score 24; DB 2; Length 33;  
Best Local Similarity 62.5%; Pred. No. 4.6e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLMGYIPA 8  
Db 15 DLMAYCEA 22

RESULT 12  
S22222  
probable lipid transfer protein - Italian stone pine (fragment)  
C:Species: Pinus pinea (Italian stone pine)  
C>Date: 19-Mar-1997 #sequence\_revision 10-Oct-1997 #text\_change 09-Jul-2004  
C:Accession: S22222  
R:Polya, G.M.; Chandra, S.; Chung, R.; Neumann, G.M.; Hoej, P.B.  
Biochim. Biophys. Acta 1120, 273-280, 1992  
A>Title: Purification and characterization of wheat and pine small basic protein substra  
A:Reference number: S21748; MUID:92247807; PMID:1576154  
C:Accession: S22222  
A:Molecule type: protein  
A:Residues: 1-36 <POL>  
A:Cross-references: UNIPROT:P26912  
A:Experimental source: seed  
C:Function:  
A>Description: involved in the phospholipid transfer  
C:Superfamily: phospholipid transfer protein  
C:Keywords: lipid transport; phospholipid

Query Match 51.1%; Score 24; DB 2; Length 36;  
Best Local Similarity 50.0%; Pred. No. 5e+02;  
Matches 3; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 3 MGYIPA 8  
Db 15 VGYLPS 20

RESULT 13  
T44523  
hypothetical protein 4S [imported] - Shigella sonnei (fragment)  
C:Species: Shigella sonnei  
C>Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T44523  
R:Chida, T.; Okamura, N.; Yoshida, Y.; Ohtani, K.; Arakawa, E.; Watanabe, H.  
submitted to the EMBL Data Library, May 1999  
A>Description: Complete DNA sequence of the O-antigen (rfb) gene cluster in Plesiomonas  
A:Reference number: Z22787  
A:Accession: T44523  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-49 <CHI>  
A:Cross-references: UNIPROT:Q9S0U2; EMBL:AB028135; PIDN:BA085074.1  
A:Experimental source: isolate HM383

Query Match 51.1%; Score 24; DB 2; Length 49;  
Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIP 7  
Db 36 GYIP 39

RESULT 14  
S38292  
30K allergen - rye (fragment)  
C:Species: Secale cereale (rye)  
C>Date: 19-May-1994 #sequence\_revision 27-Feb-1997 #text\_change 09-Jul-2004  
C:Accession: S38292  
R:Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.  
Biol. Chem. Hoppe-Seyler 374, 855-861, 1993  
A>Title: Comparison of four grass pollen species concerning their allergens of grass gro  
A:Reference number: S38288; MUID:94092339; PMID:7505598  
C:Accession: S38292  
A:Molecule type: protein  
A:Residues: 1-16 <PET>  
A:Cross-references: UNIPROT:Q7M263

Query Match 48.9%; Score 23; DB 2; Length 16;  
Best Local Similarity 80.0%; Pred. No. 3.5e+02;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GYIPA 8  
Db 4 GYAPA 8

RESULT 15  
S71871  
Nucleothione transferase (EC 2.5.1.18) class mu 10 - pig (fragments)  
A:Alternate names: glutathione S-transferase class mu 10  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 19-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 07-May-1999  
C:Accession: S71871  
R:Rouimi, P.; Anglade, P.; Debrauwer, L.; Tulliez, J.  
Biochem. J. 317, 879-884, 1996  
A>Title: Characterization of pig liver glutathione S-transferases using HPLC-electrospra  
A:Reference number: S71864; MUID:96332484; PMID:8760377  
C:Accession: S71871  
A:Molecule type: protein  
A:Residues: 1-19 <ROU>  
C:Comment: At least five species-independent classes of cytosolic glutathione transferase  
s mitochondrial form are known.  
C:Complex: dimer  
C:Function:



A;Description: catalyzes the nucleophilic conjugation of intracellular glutathione to a  
 A;Pathway: detoxification; xenobiotics metabolism  
 A;Note: increased hydrophilicity of GSH-conjugates facilitates their further metabolism  
 es of damage  
 C;Superfamily: glutathione transferase  
 C;Keywords: dimer; transferase

Query Match 48.9%; Score 23; DB 2; Length 19;  
 Best Local Similarity 60.0%; Pred. No. 4.1e+02;  
 Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLMGY 5  
 Db 2 DILGY 6

## RESULT 16

S38288  
 50K allergen - perennial ryegrass (fragment)  
 C;Species: Lolium perenne (perennial ryegrass)  
 C;Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 09-Jul-2004  
 C;Accession: S38288  
 R;Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.  
 Biol. Chem. Hoppe-Seyler 374, 855-861, 1993  
 A;Title: Comparison of four grass pollen species concerning their allergens of grass gro  
 A;Reference number: S38288; MUID:94092339; PMID:7505588  
 A;Accession: S38288  
 A;Molecule type: protein  
 A;Residues: 1-20 <PET>  
 A;Cross-references: UNIPROT:Q7M264

Query Match 48.9%; Score 23; DB 2; Length 20;  
 Best Local Similarity 80.0%; Pred. No. 4.4e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GYIPA 8  
 Db 4 GYAPA 8

## RESULT 17

S38294  
 32K allergen - common timothy (fragment)  
 C;Species: Phleum pratense (common timothy)  
 C;Date: 19-May-1994 #sequence\_revision 27-Feb-1997 #text\_change 09-Jul-2004  
 C;Accession: S38294  
 R;Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.  
 Biol. Chem. Hoppe-Seyler 374, 855-861, 1993  
 A;Title: Comparison of four grass pollen species concerning their allergens of grass gro  
 A;Reference number: S38288; MUID:94092339; PMID:7505588  
 A;Accession: S38294  
 A;Molecule type: protein  
 A;Residues: 1-20 <PET>  
 A;Cross-references: UNIPROT:Q9S8X9

Query Match 48.9%; Score 23; DB 2; Length 20;  
 Best Local Similarity 80.0%; Pred. No. 4.4e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GYIPA 8  
 Db 4 GYAPA 8

## RESULT 18

S38291  
 30K allergen - velvet grass (fragment)  
 C;Species: Holcus lanatus (velvet grass)  
 C;Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 09-Jul-2004  
 C;Accession: S38291  
 R;Petersen, A.; Schramm, G.; Becker, W.M.; Schlaak, M.  
 Biol. Chem. Hoppe-Seyler 374, 855-861, 1993  
 A;Title: Comparison of four grass pollen species concerning their allergens of grass gro

A;Reference number: S38288; MUID:94092339; PMID:7505588  
 A;Accession: S38291  
 A;Molecule type: protein  
 A;Residues: 1-20 <PET>  
 A;Cross-references: UNIPROT:Q7M262

Query Match 48.9%; Score 23; DB 2; Length 20;  
 Best Local Similarity 80.0%; Pred. No. 4.4e+02;  
 Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GYIPA 8  
 Db 4 GYTPA 8

## RESULT 19

H82216  
 hypothetical protein VCI292 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
 C;Species: Vibrio cholerae  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C;Accession: H82216  
 R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
 Chardonson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
 I, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
 Nature 406, 477-483, 2000  
 A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
 A;Reference number: A82035; MUID:20406833; PMID:10952301  
 A;Accession: H82216  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-33 <HEI>  
 A;Cross-references: UNIPROT:Q9KSG4; GB:AE004209; GB:AE003852; NID:G9655779; PIDN:AAF9445  
 A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
 C;Genetics:  
 A;Gene: VCI292  
 A;Map position: 1

Query Match 48.9%; Score 23; DB 2; Length 33;  
 Best Local Similarity 80.0%; Pred. No. 7.3e+02;  
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 4 GYIPA 8  
 Db 9 GHIPA 13

## RESULT 20

A55758  
 myosin - human (fragment)  
 C;Species: Homo sapiens (man)  
 C;Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 09-Jul-2004  
 C;Accession: A55758  
 R;Bement, W.M.; Hasson, T.; Wirth, J.A.; Cheney, R.E.; Mooseker, M.S.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 6549-6553, 1994  
 A;Title: Identification and overlapping expression of multiple unconventional myosin gen  
 A;Reference number: A55758; MUID:94294418; PMID:8022818  
 A;Accession: A55758  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-34 <RES>  
 A;Cross-references: UNIPROT:Q13871; GB:L29137; NID:G457231; PIDN:AAA20900.1; PID:G531130  
 C;Superfamily: brush border myosin heavy chain I; myosin motor domain homology

Query Match 48.9%; Score 23; DB 2; Length 34;  
 Best Local Similarity 50.0%; Pred. No. 7.5e+02;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 LMGYIPAV 9  
 Db 6 VMSYVAIV 13

## RESULT 21

A61077  
 Glue protein - ribbed mussel (fragments)  
 C;Species: Geukensia demissa (ribbed mussel)  
 C;Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 09-Jul-2004  
 C;Accession: A61077  
 R;Waite, J.H.; Hansen, D.C.; Little, K.T.  
 J. Comp. Physiol. B 159, 517-525, 1989  
 A;Title: The glue protein of ribbed mussels (Geukensia demissa): a natural adhesive with  
 A;Reference number: A61077; MUID:90110595; PMID:2481690  
 A;Accession: A61077  
 A;Molecule type: protein  
 A;Residues: 1-35 <WAI>  
 A;Cross-references: UNIPROT:Q7M3X7  
 C;Superfamily: glutenin  
 C;Keywords: hydroxylysine; hydroxyproline; tandem repeat  
 F;3,7,12,21/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status experimental  
 F;14/Modified site: 4-hydroxyproline (Pro) #status experimental  
 F;16,30,34/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) (partial) #status experimental

Query Match 48.9%; Score 23; DB 2; Length 35;  
 Best Local Similarity 75.0%; Pred. No. 7.7e+02;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GYIP 7  
 ||:|  
 Db 11 GYVP 14  
 .

RESULT 22  
 PS0339  
 immunoglobulin-binding 30K protein - Pseudomonas maltophilia (fragments)  
 C;Species: Pseudomonas maltophilia  
 C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
 C;Accession: PS0339  
 R;Grover, S.; Odell, W.D.  
 Biochem. Biophys. Res. Commun. 182, 1075-1081, 1992  
 A;Title: Partial characterization of the 30 KD Ig-binding protein from Pseudomonas maltophilia  
 A;Reference number: PS0339; MUID:92171911; PMID:1540156  
 A;Accession: PS0339  
 A;Molecule type: protein  
 A;Residues: 1-38 <GRO>  
 A;Cross-references: UNIPROT:Q7M0Q2  
 C;Comment: This cell wall protein binds the Fc region of immunoglobulin G.

Query Match 48.9%; Score 23; DB 2; Length 38;  
 Best Local Similarity 75.0%; Pred. No. 8.4e+02;  
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GYIP 7  
 ||:|  
 Db 12 GYVP 15

RESULT 23  
 E81889  
 hypothetical protein NMA1214 [imported] - Neisseria meningitidis (strain Z2491 serogroup  
 C;Species: Neisseria meningitidis  
 C;Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 09-Jul-2004  
 C;Accession: E81889  
 R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel  
 ; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,  
 Nature 404, 502-506, 2000  
 A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.  
 A;Reference number: A81775; MUID:20222556; PMID:10761919  
 A;Accession: E81889  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-38 <PAR>  
 A;Cross-references: UNIPROT:Q9JUP1; GB:AL162755; GB:AL157959; NID:G7379742; PIDN:CAB8447  
 A;Experimental source: serogroup A, strain Z2491  
 C;Genetics:  
 A;Gene: NMA1212; NMA1214

Query Match 48.9%; Score 23; DB 2; Length 38;  
 Best Local Similarity 55.8%; Pred. No. 8.4e+02;  
 Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
 : || || ||  
 Db 24 EMIIYYPVAV 32

RESULT 24  
 G95916  
 hypothetical exopeptidase protein [imported] - Sinorhizobium meliloti (strain 1021) magapla  
 C;Species: Sinorhizobium meliloti  
 C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004  
 C;Accession: G95916  
 R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
 proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
 A;Title: The complete sequence of the 1.683-kb pSymb megaplasmid from the N2-fixing endo  
 A;Reference number: A95842; MUID:21396508; PMID:11481431  
 A;Accession: G95916  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-45 <KUR>  
 A;Cross-references: UNIPROT:Q92VU9; GB:AL591985; PIDN:CAC48999.1; PID:g15140484; GSPDB:G  
 A;Experimental source: strain 1021, megaplasmid pSymb  
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, P.; Barloy-Hubler,  
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
 L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,  
 heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A;Reference number: A96039; MUID:21368234; PMID:11474104  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: SMD21026  
 A;Genome: plasmid

Query Match 48.9%; Score 23; DB 2; Length 45;  
 Best Local Similarity 33.3%; Pred. No. 1e+03;  
 Matches 2; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMGYIP 7  
 : ||:|  
 Db 22 ILGFVP 27

RESULT 25  
 I40735  
 ltua protein - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
 C;Species: Chlamydia trachomatis  
 C;Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004  
 C;Accession: I40735; E71521  
 R;Fahr, M.J.; Douglass, A.L.; Xia, W.; Hatch, T.P.  
 J. Bacteriol. 177, 4252-4260, 1995  
 A;Title: Characterization of late gene promoters of Chlamydia trachomatis.  
 A;Reference number: I40734; MUID:95362647; PMID:7543468  
 A;Accession: I40735  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-46 <RES>  
 A;Cross-references: UNIPROT:Q46403; GB:IL40822; NID:9886216; PIDN:AAA75629.1; PID:9886218  
 R;Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,  
 Science 282, 754-759, 1998  
 A;Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia tract  
 A;Reference number: A71570; MUID:9900809; PMID:9784136  
 A;Accession: E71521  
 A;Molecule type: DNA  
 A;Residues: 1-46 <ARN>  
 A;Cross-references: GB:AE001311; GB:AE001273; NID:G3328799; PIDN:AA67973.1; PID:G3328802  
 A;Experimental source: serotype D, strain UW-3/Cx  
 C;Genetics:  
 A;Gene: ltua

Query Match 48.9%; Score 23; DB 2; Length 46;  
Best Local Similarity 50.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;  
Matches 4; Conservative 2; Mismatches 0; Gaps 0;

Qy 1 DLMGYIPA 8  
| : | : |  
Db 13 DVHGYLAA 20

RESULT 26  
D81678  
conserved hypothetical protein TC0656 [imported] - Chlamydia muridarum (strain Nigg)  
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 09-Jul-2004  
C:Accession: D81678  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberger, J.F.; White, O.; Hickey,  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.  
A:Reference number: A81500; MUID:20150255; PMID:10684935  
A:Accession: D81678  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-46 <TET>  
A:Cross-references: UNIPROT:Q9PK17; GB:AE002334; GB:AE002160; NID:g7190690; PIDN:AAF3948  
A:Experimental source: strain Nigg (MoPn)  
C:Genetics:  
A:Gene: TC0656

Query Match 48.9%; Score 23; DB 2; Length 46;  
Best Local Similarity 50.0%; Pred. No. 1e+03; 2; Indels 0; Gaps 0;  
Matches 4; Conservative 2; Mismatches 0; Gaps 0;

Qy 1 DLMGYIPA 8  
| : | : |  
Db 13 DVHGYLAA 20

RESULT 27  
PH1312  
Ig heavy chain DJ region (clone C68-105) - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-May-1999  
C:Accession: PH1312  
R:Wasserman, R.; Gallili, N.; Ito, Y.; Reichard, B.A.; Shane, S.; Rovera, G.  
J. Exp. Med. 176, 1577-1581, 1992  
A:Title: Predominance of fetal type DJH joining in young children with B precursor lymphoma.  
A:Reference number: PH1302; MUID:93094761; PMID:1460419  
A:Accession: PH1312  
A:Molecule type: DNA  
A:Residues: 1-17 <WAS>  
C:Keywords: heterotetramer; immunoglobulin

Query Match 46.8%; Score 22; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 5.9e+02; 0; Indels 0; Gaps 0;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 YIPA 8  
| : | : |  
Db 7 YIPA 10

RESULT 28  
A05332  
serum amyloid P-component - plaice (fragment)  
N:Alternate names: SAP  
C:Species: Pleuronectes platessa (plaice)  
C:Date: 05-Jun-1987 #sequence\_revision 05-Jun-1987 #text\_change 09-Jul-2004  
C:Accession: A05332  
R:Pepys, M.B.; de Beer, F.C.; Milstein, C.P.; March, J.P.; Feinstein, A.; Butress, N.; C  
Biochim. Biophys. Acta 704, 123-133, 1982  
A:Reference number: A90642; MUID:82232106; PMID:7093286

A:Contents: tentative sequence  
A:Accession: A05332  
A:Molecule type: protein  
A:Residues: 1-20 <PEP>  
A:Cross-references: UNIPROT:P20677  
C:Superfamily: C-reactive protein  
C:Keywords: amyloid; glycoprotein; pentraxin; plasma

Query Match 46.8%; Score 22; DB 2; Length 20;  
Best Local Similarity 66.7%; Pred. No. 6.9e+02; 1; Indels 0; Gaps 0;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYI 6  
| : | : |  
Db 4 DLMGV 9

RESULT 29  
S06149  
photosystem I chain III - barley (fragment)  
C:Species: Hordeum vulgare (barley)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 18-Jun-1993  
C:Accession: S06149  
R:Anandan, S.; Vainstein, A.; Thornber, J.P.  
FEBS Lett. 256, 150-154, 1989  
A:Title: Correlation of some published amino acid sequences for photosystem I polypeptide  
A:Reference number: S06149; MUID:90033290; PMID:2680596  
A:Accession: S06149  
A:Molecule type: protein  
A:Residues: 1-20 <ANA>  
A:Note: the authors identified this protein as light-harvesting complex I 17K protein  
C:Keywords: chloroplast; photosynthesis; photosystem I; thylakoid

Query Match 46.8%; Score 22; DB 2; Length 20;  
Best Local Similarity 50.0%; Pred. No. 6.9e+02; 3; Indels 0; Gaps 0;  
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 DLMGYIPA 8  
| : | : |  
Db 1 DIAGLTPA 8

RESULT 30  
S21772  
lipoxigenase (EC 1.13.11.12) 1 - barley (fragment)  
C:Species: Hordeum vulgare (barley)  
C:Date: 19-Mar-1997 #sequence\_revision 24-Mar-1999 #text\_change 24-Mar-1999  
C:Accession: S21772  
R:Doderer, A.; Kokkelink, I.; van der Veen, S.; Valk, B.E.; Schram, A.W.; Douma, A.C.  
Biochim. Biophys. Acta 1120, 97-104, 1992  
A:Title: Purification and characterization of two lipoxigenase isoenzymes from germinating  
A:Reference number: S21772; MUID:92207997; PMID:1554746  
A:Accession: S21772  
A:Molecule type: protein  
A:Residues: 1-21 <DOD>  
A:Experimental source: var. Triumph, seed  
C:Function:  
A:Description: catalyzes the peroxidation of polyunsaturated fatty acids to their corres  
C:Superfamily: lipoxigenase  
C:Keywords: monomer; oxidoreductase; seed

Query Match 46.8%; Score 22; DB 2; Length 21;  
Best Local Similarity 66.7%; Pred. No. 7.3e+02; 1; Indels 0; Gaps 0;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 GYIPAV 9  
| : | : |  
Db 5 GILPAV 10

RESULT 31  
E27011  
MHC class II-associated invariant chain - human



A;Title: Sequences of six genes and several open reading frames in the kinetoplast maxid  
A;Reference number: A22848; MUID:85079995; PMID:6096360

A;Accession: A30010  
A;Reference number: A22848; MWID:850/55555; PMID:8056300  
A;Molecule type: DNA  
A;Residues: 1-42 <DEL>  
A;Cross-references: GB:M10126  
C;Genetics:  
A;Genome: mitochondrion  
A;Genetic code: SGC6  
C;Superfamily: cytochrome-c oxidase chain II, mitochondrial type; cytochrome-c oxidase c  
C;Keywords: copper; electron transfer; membrane-associated complex; mitochondrial inner

Query Match	46.8%	Score 22;	DB 2;	Length 42;
Best Local Similarity	50.0%	Pred. No. 1.5e+03;		
Matches	4;	Conservative	2;	Mismatches 2;
				Indels 0;
				Gaps 0;

Qy 2 LMGYIPAV 9  
| | | | |  
Dy 31 LMGEMPV 38

RESULT 37  
PS0359  
hypothetical protein (VP16 3' region) - human herpesvirus 2 (fragment)  
C:Species: human herpesvirus 2  
C.Date: 17-Aug-1992 #sequence revision 17-Aug-1992 #text change 09-Jul-2004

C: Superfamily: herpesvirus 70-5K alpha trans-inducing protein  
A: Experimental source: strain HG52  
A: Cross-references: UNIPROT:P89467; GB:M60050  
A: Residues: 1-43 <CRE>  
A: Molecule type: DNA  
A: Accession: P80359  
A: Reference number: JS0689; MUID:91365250; PMID:1653757  
A: Title: Nucleotide and deduced amino acid sequences of the gene encoding virion protein  
Gene 103, 235-238, 1991  
R: Cress, A.; Triesenberg, S.J.  
C: Accession: F80355

Query Match 46.8%; Score 22; DB 2; Length 43;  
Best Local Similarity 100.0%; Pred. No. 1.5e+03;  
Matches 4.; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LMGY 5  
40 LMGY 43

RESULT 38  
C35156  
dihydrolipoamide dehydrogenase (EC 1.8.1.4) - Peptostreptococcus glycinophilus (fragment)  
C:Species: Peptostreptococcus glycinophilus  
C:Date: 10-Aug-1990 #sequence\_revision 10-Aug-1990 #text\_change 09-Jul-2004  
C:Accession: C35156  
J:Dietrichs, D.; Meyer, M.; Schmidt, B.; Andreessen, J.R.  
J. Bacteriol. 172, 2088-2095, 1990  
A:Title: Purification of NAD(P)-dependent electron-transferring flavoproteins and N-termi  
A:Reference number: A35156; MUID: 90202731; PMID:2318809

Query Match 46.8%; Score 22; DB 2; Length 44;  
Best Local Similarity 50.0%; Pred. No. 1.5e+03;  
Matches 1. Conservative 2: Mismatches 1; Indels 0  
Gaps 0

QY 4 GYIPAV 9  
|||:  
Db 15 GYVAAT 20

## RESULT 39

B35156  
 dihydrolipoamide dehydrogenase (EC 1.8.1.4) - Clostridium sporogenes (fragment)  
 C:Species: Clostridium sporogenes  
 C:Date: 10-Aug-1990 #sequence\_revision 10-Aug-1990 #text\_change 09-Jul-2004  
 C:Accession: B35156  
 R:Dieterichs, D.; Meyer, M.; Schmidt, B.; Andresen, J.R.

J. Bacteriol. 172: 2088-2095, 1990.  
 A: Title: Purification of NADPH-dependent electron-transferring flavoproteins and N-termin  
 A: Reference number: A35156; MUID: 90202731; PMID: 2318809  
 A: Accession: B35156  
 A: Status: preliminary  
 A: Molecule type: protein  
 A: Residues: 1-45 <DIE>  
 A: Cross-references: UNIPROT:Q7M0Y4  
 C: Superfamily: dihydroliipoamide dehydrogenase; dihydroliipoamide dehydrogenase homology  
 C: Keywords: FAD; flavoprotein; homodimer; liipoamide; NAD; oxidoreductase  
 C: 3-31/Region: beta-alpha-beta FAD nucleotide-binding fold

Query Match 46.8%; Score 22; DB 2; Length 45;  
Best Local Similarity 50.0%; Pred. No. 1.6e+03;  
Matches 3: Conservative 2; Mismatches 1; Indels

Qy 4 GYIPAV 9  
||:|:  
Db 13 GYVAAI 18

## RESULT 40

N:Alternate names: Desulfovibrio vulgaris (strain Miyazaki) (fragments)  
 C:Species: Desulfovibrio vulgaris  
 C:Date: 19-Mar-1997 #sequence\_revision 18-Jun-1999 #text\_change 09-Jul-2004  
 C:Accession: S16924  
 R:Tasaka, C.; Ogata, M.; Yagi, T.; Tsugita, A.  
 Protein Seq. Data Anal. 4, 25-27, 1991  
 A:Title: Partial sequences of high-molecular-weight cytochrome c isolated f  
 A:Reference number: S16924: MUID:92020808: PMID:16564429

A:Accession: J06924  
A:Molecule type: protein  
A:Residues: 1-30;31-48 <PRO>  
A:Cross-references: UNIPROT:Q7M0Z0  
A:Experimental source: strain Miyazaki  
C:Superfamily: cytochrome cc3; cytochrome c3 homolog  
C:Keywords: blocked amino end; chromoprotein; heme; iron

Query Match	46.8%	Score 22;	DB 2;	Length 48;
Best Local Similarity	57.1%;	Pred. NO. 1.7e+03;		
Matches 4:	Conservative	2:	Mismatches 1;	Indels 0;
				Gaps 0;

QY 3 MGYIPAV 9  
:|:|  
30 VGTLPVAV 36

Search completed: August 23, 2005, 13:36:34  
Job time : 43 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2005, 13:35:09 ; Search time 158 Seconds  
(without alignments)

22.306 Million cell updates/sec

Title: US-10-770-117-1

Perfect score: 47

Sequence: 1 DLNGYIPAV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 507806

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11A\_PUBCOMB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
21: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	100.0	9	14	US-10-114-823B-3
2	47	100.0	9	18	US-10-770-117-1
3	47	100.0	9	20	US-11-082-595-3
4	42	89.4	9	20	US-08-854-825-54
5	42	89.4	9	9	US-09-894-018-212
6	42	89.4	9	14	US-10-128-711-51
7	42	89.4	9	14	US-10-128-711-130
8	42	89.4	9	14	US-10-114-823B-20
9	42	89.4	9	15	US-10-371-525-255
10	42	89.4	9	15	US-10-371-069-255
11	42	89.4	9	15	US-10-371-645-255
12	42	89.4	9	15	US-10-371-260-255
13	42	89.4	9	15	US-10-367-580-74
14	42	89.4	9	15	US-10-367-593-74
15	42	89.4	9	15	US-10-367-594-74
16	42	89.4	9	15	US-10-367-654-74
17	42	89.4	9	15	US-10-367-658-74
18	42	89.4	9	15	US-10-367-668-74
19	42	89.4	9	16	US-10-440-390-2
20	42	89.4	9	16	US-10-367-674-74
21	42	89.4	9	16	US-10-777-053-199
22	42	89.4	9	16	US-10-777-053-200
23	42	89.4	9	16	US-10-777-053-426
24	42	89.4	9	16	US-10-837-217-199
25	42	89.4	9	16	US-10-837-217-200
26	42	89.4	9	16	US-10-837-217-426
27	42	89.4	9	16	US-10-474-960A-212
28	42	89.4	9	18	US-10-481-696-19
29	42	89.4	9	18	US-10-770-117-3
30	42	89.4	9	20	US-11-082-595-20
31	42	89.4	10	8	US-08-854-825-1
32	42	89.4	10	16	US-10-777-053-109
33	42	89.4	10	16	US-10-837-217-109
34	42	89.4	10	17	US-10-931-566-26
35	42	89.4	15	13	US-10-103-395-172
36	42	89.4	15	14	US-10-268-561-14
37	42	89.4	15	14	US-10-268-569-14
38	42	89.4	20	9	US-09-894-018-4
39	42	89.4	20	14	US-10-114-823B-2
40	42	89.4	20	15	US-10-651-165-10
41	42	89.4	20	16	US-10-474-960A-4
42	42	89.4	20	20	US-11-082-595-2
43	42	89.4	21	9	US-09-894-018-3
44	42	89.4	22	14	US-10-296-558-7
45	42	89.4	22	14	US-10-474-960A-3
46	42	89.4	24	15	US-10-651-165-216
47	42	89.4	25	10	US-09-891-983A-31
48	42	89.4	25	10	US-09-891-983A-32
49	42	89.4	25	14	US-10-173-480-58
50	42	89.4	25	14	US-10-173-480-59
51	42	89.4	25	16	US-10-753-910-58
52	42	89.4	25	16	US-10-753-910-59
53	42	89.4	30	15	US-10-296-734-424
54	42	89.4	37	15	US-10-651-165-281
55	39	83.0	8	9	US-09-989-621-10
56	38	80.9	24	15	US-10-651-165-285
57	36	76.6	8	15	US-10-651-165-69
58	36	76.6	9	15	US-10-651-165-163
59	36	76.6	10	15	US-10-367-580-75
60	36	76.6	10	15	US-10-367-593-75
61	36	76.6	10	15	US-10-367-594-75
62	36	76.6	10	15	US-10-367-654-75
63	36	76.6	10	15	US-10-367-658-75
64	36	76.6	10	15	US-10-367-668-75
65	36	76.6	10	16	US-10-367-674-75
66	36	76.6	10	16	US-10-777-053-108
67	36	76.6	10	16	US-10-837-217-108
68	36	76.6	20	15	US-10-651-165-11
69	36	76.6	44	15	US-10-651-165-67
70	33	70.2	28	14	US-10-109-670-45
71	32	68.1	7	15	US-10-651-165-70
72	32	68.1	30	15	US-10-296-734-426
73	31	66.0	30	14	US-09-764-891-3423
74	31	66.0	30	14	US-10-091-573-301
75	31	66.0	41	16	US-10-425-115-291449
76	31	66.0	41	16	US-10-425-115-291482
77	31	66.0	43	14	US-10-411-224-171
78	31	66.0	43	15	US-10-047-021-171
79	29	61.7	35	16	US-10-425-115-220432
80	29	61.7	38	14	US-10-029-386-31676
81	29	61.7	40	16	US-10-425-115-270620
82	29	61.7	43	16	US-10-425-115-256302
83	28	59.6	9	17	US-10-654-601-2277
84	28	59.6	10	9	US-09-766-412-20

Sequence 255, App  
Sequence 74, Appl  
Sequence 74, Appl  
Sequence 74, Appl  
Sequence 74, Appl  
Sequence 74, Appl  
Sequence 74, Appl  
Sequence 2, Appl  
Sequence 74, Appl  
Sequence 199, App  
Sequence 200, App  
Sequence 426, App  
Sequence 200, App  
Sequence 426, App  
Sequence 212, App  
Sequence 19, Appl  
Sequence 3, Appl  
Sequence 20, Appl  
Sequence 1, Appl  
Sequence 109, App  
Sequence 109, App  
Sequence 26, Appl  
Sequence 172, App  
Sequence 14, Appl  
Sequence 14, Appl  
Sequence 4, Appl  
Sequence 2, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 2, Appl  
Sequence 3, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 216, App  
Sequence 31, Appl  
Sequence 32, Appl  
Sequence 58, Appl  
Sequence 58, Appl  
Sequence 59, Appl  
Sequence 59, Appl  
Sequence 424, App  
Sequence 281, App  
Sequence 10, Appl  
Sequence 285, App  
Sequence 69, Appl  
Sequence 163, App  
Sequence 75, Appl  
Sequence 75, Appl  
Sequence 75, Appl  
Sequence 75, Appl  
Sequence 75, Appl  
Sequence 75, Appl  
Sequence 108, App  
Sequence 108, App  
Sequence 11, Appl  
Sequence 67, Appl  
Sequence 45, Appl  
Sequence 70, Appl  
Sequence 426, App  
Sequence 3423, App  
Sequence 301, App  
Sequence 291449, App  
Sequence 291482, App  
Sequence 171, App  
Sequence 220432, App  
Sequence 31676, A  
Sequence 270620, App  
Sequence 256302, App  
Sequence 2277, App  
Sequence 20, Appl

Sequence 232, App  
 Sequence 284, App  
 Sequence 284, App  
 Sequence 284, App  
 Sequence 284, App  
 Sequence 232, App  
 Sequence 9, Appl  
 Sequence 10, Appl  
 Sequence 1230, Ap  
 Sequence 7, Appl  
 Sequence 8, Appl  
 Sequence 380, App  
 Sequence 414, App  
 Sequence 127, App  
 Sequence 6, Appl  
 Sequence 271139,

## ALIGNMENTS

RESULT 1  
 US-10-114-823B-3  
 ; Sequence 3, Application US/10114823B  
 ; Publication No. US20030099663A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEITMANN, JULIA-KRISTINA  
 ; APPLICANT: MATTNER, FRANK  
 ; APPLICANT: BUSCHLE, MICHAEL  
 ; APPLICANT: MELLING, JACK  
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION COMPRISING AN ANTIGEN  
 ; FILE REFERENCE: SONN:0160US  
 ; CURRENT APPLICATION NUMBER: US/10/114,823B  
 ; CURRENT FILING DATE: 2002-04-01  
 ; PRIOR APPLICATION NUMBER: PCT/EP00/09657  
 ; PRIOR FILING DATE: 2000-10-02  
 ; PRIOR APPLICATION NUMBER: A-1680/99  
 ; PRIOR FILING DATE: 1999-10-01  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 ; OTHER INFORMATION: Peptide  
 US-10-114-823B-3

Query Match 100.0%; Score 47; DB 14; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
 |||||

Db 1 DLMGYIPAV 9

RESULT 2  
 US-10-770-117-1  
 ; Sequence 1, Application US/10770117  
 ; Publication No. US20050129705A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: The Government of the United States, as represented by the  
 ; APPLICANT: Secretary, Dept. of  
 ; APPLICANT: Health and Human Services  
 ; APPLICANT: Berzofsky, Jay A.  
 ; APPLICANT: Sarobe, Pablo  
 ; APPLICANT: Pendleton, C. David  
 ; APPLICANT: Feinstein, Stephen M.  
 ; APPLICANT: Arichi, Tatsumi  
 ; APPLICANT: Major, Marian E.  
 ; TITLE OF INVENTION: MODIFIED HCV PEPTIDE VACCINES

FILE REFERENCE: 14014.0347/P  
 ; CURRENT APPLICATION NUMBER: US/10/770,117  
 ; CURRENT FILING DATE: 2004-02-02  
 ; PRIOR APPLICATION NUMBER: US/09/763,260  
 ; PRIOR FILING DATE: 2001-10-19  
 ; PRIOR APPLICATION NUMBER: 60/097,446  
 ; PRIOR FILING DATE: 1998-08-21  
 ; NUMBER OF SEQ ID NOS: 8  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Note =  
 ; OTHER INFORMATION: synthetic construct  
 US-10-770-117-1

Query Match 100.0%; Score 47; DB 18; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
 |||||

Db 1 DLMGYIPAV 9

RESULT 3  
 US-11-082-595-3  
 ; Sequence 3, Application US/11082595  
 ; Publication No. US20050163797A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FLEITMANN, JULIA-KRISTINA  
 ; APPLICANT: MATTNER, FRANK  
 ; APPLICANT: BUSCHLE, MICHAEL  
 ; APPLICANT: MELLING, JACK  
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION COMPRISING AN ANTIGEN  
 ; FILE REFERENCE: SONN:0160US  
 ; CURRENT APPLICATION NUMBER: US/11/082,595  
 ; CURRENT FILING DATE: 2005-03-17  
 ; PRIOR APPLICATION NUMBER: US/10/114,823  
 ; PRIOR FILING DATE: 2002-04-01  
 ; PRIOR APPLICATION NUMBER: PCT/EP00/09657  
 ; PRIOR FILING DATE: 2000-10-02  
 ; PRIOR APPLICATION NUMBER: A-1680/99  
 ; PRIOR FILING DATE: 1999-10-01  
 ; NUMBER OF SEQ ID NOS: 24  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 9  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
 ; OTHER INFORMATION: Peptide  
 US-11-082-595-3

Query Match 100.0%; Score 47; DB 20; Length 9;  
 Best Local Similarity 100.0%; Pred. No. 1.6e+06;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
 |||||

Db 1 DLMGYIPAV 9

RESULT 4  
 US-08-854-825-54  
 ; Sequence 54, Application US/08854825  
 ; Publication No. US20020115061A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chisari, Francis V.  
 ; APPLICANT: Cerny, Andreas



```

; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Leydig, Voit & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,825
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Silver, Donald J.
; REGISTRATION NUMBER: 37552
; REFERENCE/DOCKET NUMBER: 61230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-854-825-54

Query Match 89.4%; Score 42; DB 8; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9
   |||||
Db 1 DLMGYIPLV 9

RESULT 5
US-09-894-018-212
; Sequence 212, Application US/09894018
; Patent No. US20020119127A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Chestnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denisw
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: METHODS AND SYSTEM FOR OPTIMIZING
; TITLE OF INVENTION: MINIGENES AND PEPTIDES THEREBY
; FILE REFERENCE: 39963-20033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: PCT/US00/35568
; PRIOR FILING DATE: 2000-12-28
; PRIOR APPLICATION NUMBER: US 60/173,390
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 212
; LENGTH: 9
; TYPE: PRT

```

```

; ORGANISM: Transgenic mouse
; US-09-894-018-212

Query Match 89.4%; Score 42; DB 9; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9
   |||||
Db 1 DLMGYIPLV 9

RESULT 6
US-10-128-711-51
; Sequence 51, Application US/10128711
; Publication No. US20030099634A1
; GENERAL INFORMATION:
; APPLICANT: VITIELLO, Maria A.
; CHESTNUT, Robert W.
; SETTE, Alessandro D.
; CELIS, Esteban
; GRAY, Howard
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/128,711
; FILING DATE: 22-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/197,484
; FILING DATE: 16-FEB-1994
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (206) 623-6793
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-10-128-711-51

Query Match 89.4%; Score 42; DB 14; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

Qy 1 DLMGYIPAV 9  
| | | | | | |  
Db 1 DLMGYIPLV 9

## RESULT 7

US-10-128-711-130  
; Sequence 130, Application US/10128711  
; Publication No. US20030099634A1  
; GENERAL INFORMATION:  
; APPLICANT: VITIELLO, Maria A.  
; CHESTNUT, Robert W.  
; SETTE, Alessandro D.  
; CELIS, Esteban  
; GRAY, Howard  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/128,711  
; FILING DATE: 22-Apr-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/197,484  
; FILING DATE: 16-FEB-1994  
; APPLICATION NUMBER: US 07/935,811  
; FILING DATE: 26-AUG-1992  
; APPLICATION NUMBER: US 07/874,491  
; FILING DATE: 27-APR-1992  
; APPLICATION NUMBER: US 07/827,682  
; FILING DATE: 29-JAN-1992  
; APPLICATION NUMBER: US 07/749,568  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; REFERENCE/DOCKET NUMBER: 14137-26-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 467-9600  
; TELEFAX: (206) 623-6793  
; INFORMATION FOR SEQ ID NO: 130:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 130:  
US-10-128-711-130

Query Match 89.4%; Score 42; DB 14; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.6e+06;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | |  
Db 1 DLMGYIPLV 9

## RESULT 8

US-10-114-823B-20

; Sequence 20, Application US/10114823B  
; Publication No. US20030099663A1  
; GENERAL INFORMATION:  
; APPLICANT: FLEITMANN, JULIA-KRISTINA  
; APPLICANT: MATTNER, FRANK  
; APPLICANT: BUSCHLE, MICHAEL  
; APPLICANT: MELLING, JACK  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION COMPRISING AN ANTIGEN  
; FILE REFERENCE: SONN:016US  
; CURRENT APPLICATION NUMBER: US/10/114,823B  
; CURRENT FILING DATE: 2002-04-01  
; PRIOR APPLICATION NUMBER: PCT/EP00/09657  
; PRIOR FILING DATE: 2000-10-02  
; PRIOR APPLICATION NUMBER: A-1680/99  
; PRIOR FILING DATE: 1999-10-01  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Peptide  
US-10-114-823B-20

Query Match 89.4%; Score 42; DB 14; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.6e+06;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | |  
Db 1 DLMGYIPLV 9

## RESULT 9

US-10-371-525-255  
; Sequence 255, Application US/10371525  
; Publication No. US20030203869A1  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; IMMUNE RESPONSE AND METHODS OF USING THE SAME  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/10/371,525  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/311,784  
; PRIOR FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 255  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HCV Core 132 (peptide 1013.1002)  
US-10-371-525-255

Query Match 89.4%; Score 42; DB 15; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.6e+06;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | |  
Db 1 DLMGYIPLV 9

```
RESULT 10
US-10-371-069-255
; Sequence 255, Application US/10371069
; Publication No. US20030216342A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.10
; CURRENT APPLICATION NUMBER: US/10/371.069
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 255
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV Core 132 (peptide 1013.1002)
US-10-371-069-255

Query Match      89.4%; Score 42; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      1 DLMGYIPLV 9

RESULT 11
US-10-371-645-255
; Sequence 255, Application US/10371645
; Publication No. US20030216343A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.11
; CURRENT APPLICATION NUMBER: US/10/371.645
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 255
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV Core 132 (peptide 1013.1002)
```

```
US-10-371-645-255

Query Match      89.4%; Score 42; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      1 DLMGYIPLV 9

RESULT 12
US-10-371-260-255
; Sequence 255, Application US/10371260
; Publication No. US20030220285A1
; GENERAL INFORMATION:
; APPLICANT: EPIMMUNE Inc.
; APPLICANT: Fikes, John D.
; APPLICANT: Hermanson, Gary G.
; APPLICANT: Sette, Alessandro
; APPLICANT: Ishioka, Glenn Y.
; APPLICANT: Livingston, Brian
; APPLICANT: Chesnut, Robert W.
; APPLICANT: Epimmune Inc.
; TITLE OF INVENTION: Expression Vectors for Stimulating an
; FILE OF INVENTION: Immune Response and Methods of Using the Same
; FILE REFERENCE: 39963-20022.13
; CURRENT APPLICATION NUMBER: US/10/371.260
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/078,904
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: US 60/085,751
; PRIOR FILING DATE: 1998-05-15
; NUMBER OF SEQ ID NOS: 463
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 255
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV Core 132 (peptide 1013.1002)
US-10-371-260-255

Query Match      89.4%; Score 42; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      1 DLMGYIPLV 9

RESULT 13
US-10-367-580-74
; Sequence 74, Application US/10367580
; Publication No. US20040071720A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Takechi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/461061
; CURRENT APPLICATION NUMBER: US/10/367.580
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 09/794,832
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 09/011,645
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
; PRIOR FILING DATE: 1996-08-16
```

```

; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 74
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-367-590-74

```

```

Query Match      89.4%; Score 42; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 DLMGYIPAV 9
   |||||
Db 1 DLMGYIPLV 9

```

## RESULT 14

```

US-10-367-593-74
; Sequence 74, Application US/10367593
; Publication No. US20040071721A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Takechi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/461012
; CURRENT APPLICATION NUMBER: US/10367,593
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 09/011,645
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
; PRIOR FILING DATE: 1996-08-16
; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 74
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-367-593-74

```

```

Query Match      89.4%; Score 42; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 DLMGYIPAV 9
   |||||
Db 1 DLMGYIPLV 9

```

## RESULT 15

```

US-10-367-594-74

```

```

; Sequence 74, Application US/10367594
; Publication No. US20040071721A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.

```

```

; APPLICANT: Houghton, Alan
; APPLICANT: Takechi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/461041
; CURRENT APPLICATION NUMBER: US/10/367,594
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 09/680,806
; PRIOR FILING DATE: 2000-10-05
; PRIOR APPLICATION NUMBER: US 09/011,645
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
; PRIOR FILING DATE: 1996-08-16
; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 74
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-367-594-74

```

```

Query Match      89.4%; Score 42; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 DLMGYIPAV 9
   |||||
Db 1 DLMGYIPLV 9

```

## RESULT 16

```

US-10-367-654-74

```

```

; Sequence 74, Application US/10367654
; Publication No. US20040071723A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Takechi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/461032
; CURRENT APPLICATION NUMBER: US/10/367,654
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 10/171,734
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/636,295
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: US 09/011,645
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
; PRIOR FILING DATE: 1996-08-16
; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 74
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-367-654-74

```

```
Query Match      89.4%; Score 42; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      1 DLMGYIPLV 9

RESULT 17
US-10-367-658-74
; Sequence 74, Application US/10367658
; Publication No. US20040071724A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Takechi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/461051
; CURRENT APPLICATION NUMBER: US/10/367,658
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 09/794,529
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 09/011,645
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
; PRIOR FILING DATE: 1996-08-16
; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 74
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-367-658-74

Query Match      89.4%; Score 42; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      1 DLMGYIPLV 9

RESULT 18
US-10-367-668-74
; Sequence 74, Application US/10367668
; Publication No. US20040071725A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Takechi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/461072
; CURRENT APPLICATION NUMBER: US/10/367,668
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 09/794,517
; PRIOR FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: US 09/011,645
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
```

```
Query Match      89.4%; Score 42; DB 15; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      1 DLMGYIPLV 9

RESULT 19
US-10-440-390-2
; Sequence 2, Application US/10440390
; Publication No. US20040105868A1
; GENERAL INFORMATION:
; APPLICANT: Lauer et al.
; TITLE OF INVENTION: Epitopes of Hepatitis C Virus
; FILE REFERENCE: 24028-011
; CURRENT APPLICATION NUMBER: US/10/440,390
; CURRENT FILING DATE: 2003-05-16
; PRIOR APPLICATION NUMBER: 60/381,273
; PRIOR FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; OTHER INFORMATION: synthetic peptide
US-10-440-390-2

Query Match      89.4%; Score 42; DB 16; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      1 DLMGYIPLV 9

RESULT 20
US-10-367-674-74
; Sequence 74, Application US/10367674
; Publication No. US20040127684A1
; GENERAL INFORMATION:
; APPLICANT: Rothman, James E.
; APPLICANT: Hartl, F. Ulrich
; APPLICANT: Hoe, Mee H.
; APPLICANT: Houghton, Alan
; APPLICANT: Takechi, Yoshizumi
; APPLICANT: Mayhew, Mark
; TITLE OF INVENTION: Heat Shock Protein-Based Vaccines and Immunotherapies
; FILE REFERENCE: 11746/4610211
; CURRENT APPLICATION NUMBER: US/10/367,674
; CURRENT FILING DATE: 2003-02-14
; PRIOR APPLICATION NUMBER: US 10/170,738
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/552,868
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 09/011,645
```

```
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: PCT/US96/13363
; PRIOR FILING DATE: 1996-08-16
; PRIOR APPLICATION NUMBER: US 60/002,490
; PRIOR FILING DATE: 1995-08-18
; PRIOR APPLICATION NUMBER: US 60/002,479
; PRIOR FILING DATE: 1995-08-18
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: WordPerfect 8.0 for Windows
; SEQ ID NO 74
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-367-674-74

Query Match      89.4%; Score 42; DB 16; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
      |||||
Db      1 DLMGYIPLV 9

RESULT 21
US-10-777-053-199
; Sequence 199, Application US/10777053
; Publication No. US20040132088A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; TITLE OF INVENTION: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MANNK.022C1
; CURRENT APPLICATION NUMBER: US/10/777,053
; PRIOR FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 199
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Human Cytomegalovirus
US-10-777-053-199

Query Match      89.4%; Score 42; DB 16; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
      |||||
Db      1 DLMGYIPLV 9

RESULT 22
US-10-777-053-200
; Sequence 200, Application US/10777053
; Publication No. US20040132088A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; TITLE OF INVENTION: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MANNK.022C1
; CURRENT APPLICATION NUMBER: US/10/777,053
; PRIOR FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 199
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-10-777-053-426

Query Match      89.4%; Score 42; DB 16; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
      |||||
Db      1 DLMGYIPLV 9

RESULT 23
US-10-777-053-426
; Sequence 426, Application US/10777053
; Publication No. US20040132088A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; TITLE OF INVENTION: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MANNK.022C1
; CURRENT APPLICATION NUMBER: US/10/777,053
; PRIOR FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 426
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-10-777-053-426

Query Match      89.4%; Score 42; DB 16; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
      |||||
Db      1 DLMGYIPLV 9

RESULT 24
US-10-837-217-199
; Sequence 199, Application US/10837217
; Publication No. US20040203051A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; TITLE OF INVENTION: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MANNK.022C2
; CURRENT APPLICATION NUMBER: US/10/837,217
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 10/292,413
```

```

; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 199
; TYPE: PRT
; ORGANISM: Human Cytomegalovirus
US-10-837-217-199

Query Match      89.4%; Score 42; DB 16; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DLMGYIPAV 9
DB      1 DLMGYIPLV 9

RESULT 25
US-10-837-217-200
; Sequence 200, Application US/10837217
; Publication No. US20040203051A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; TITLE OF INVENTION: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MANNK.022C2
; CURRENT APPLICATION NUMBER: US/10/837,217
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 200
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-10-837-217-200

Query Match      89.4%; Score 42; DB 16; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DLMGYIPAV 9
DB      1 DLMGYIPLV 9

RESULT 26
US-10-837-217-426
; Sequence 426, Application US/10837217
; Publication No. US20040203051A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; TITLE OF INVENTION: TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MANNK.022C2
; CURRENT APPLICATION NUMBER: US/10/837,217
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 426
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-10-837-217-426

Query Match      89.4%; Score 42; DB 16; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DLMGYIPAV 9
DB      1 DLMGYIPLV 9

RESULT 27
US-10-474-960A-212
; Sequence 212, Application US/10474960A
; Publication No. US20040248113A1
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Chesnut, Robert
; APPLICANT: Livingston, Brian
; APPLICANT: Baker, Denise
; APPLICANT: Newman, Mark
; APPLICANT: Brown, David
; TITLE OF INVENTION: Methods and System for Optimizing Multi-epitope Nucleic
; TITLE OF INVENTION: Acid Constructs and Peptides Encoded Thereby
; FILE REFERENCE: 2060.0320004
; CURRENT APPLICATION NUMBER: US/10/474,960A
; CURRENT FILING DATE: 2003-10-16
; PRIOR APPLICATION NUMBER: PCT/US02/09877
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 09/894,018
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/284,221
; PRIOR FILING DATE: 2001-04-16
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 212
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Sequence for epitope
US-10-474-960A-212

Query Match      89.4%; Score 42; DB 16; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 DLMGYIPAV 9
DB      1 DLMGYIPLV 9

RESULT 28
US-10-481-696-19
; Sequence 19, Application US/10481696
; Publication No. US20050129617A1
; GENERAL INFORMATION:
; APPLICANT: THE UNIVERSITY OF BRITISH COLUMBIA
; TITLE OF INVENTION: TYPE 1 DIABETES DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 80021-365
; CURRENT APPLICATION NUMBER: US/10/481,696
; CURRENT FILING DATE: 2003-12-22
; PRIOR APPLICATION NUMBER: US 60/299,754
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19

```

```
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV peptide epitope
US-10-481-696-19

Query Match      89.4%; Score 42; DB 18; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9
Db 1 DLMGYIPLV 9
|||||||

RESULT 29
US-10-770-117-3
; Sequence 3, Application US/10770117
; Publication No. US20050129705A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States, as represented by the
; APPLICANT: Secretary, Dept. of
; APPLICANT: Health and Human Services
; APPLICANT: Berzofsky, Jay A.
; APPLICANT: Sarobe, Pablo
; APPLICANT: Pendleton, C. David
; APPLICANT: Feinstein, Stephen M.
; APPLICANT: Arichi, Tatsumi
; APPLICANT: Major, Marian E.
; TITLE OF INVENTION: MODIFIED HCV PEPTIDE VACCINES
; FILE REFERENCE: 14014.0347/P
; CURRENT APPLICATION NUMBER: US/10/770,117
; CURRENT FILING DATE: 2004-02-02
; PRIOR FILING DATE: 2004-02-02
; PRIOR FILING DATE: 2001-10-19
; PRIOR FILING DATE: 2001-10-19
; PRIOR FILING DATE: 2001-10-19
; PRIOR FILING DATE: 1998-08-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: /Note =
; OTHER INFORMATION: synthetic construct
US-10-770-117-3

Query Match      89.4%; Score 42; DB 18; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9
Db 1 DLMGYIPLV 9
|||||||

RESULT 30
US-11-082-595-20
; Sequence 20, Application US/11082595
; Publication No. US20050163797A1
; GENERAL INFORMATION:
; APPLICANT: FLEITMANN, JULIA-KRISTINA
; APPLICANT: MATTERN, FRANK
; APPLICANT: BUSCHLE, MICHAEL
; APPLICANT: MELLING, JACK
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION COMPRISING AN ANTIGEN
; FILE REFERENCE: SONN:0160US
; CURRENT APPLICATION NUMBER: US/11/082,595
; CURRENT FILING DATE: 2005-03-17
; PRIOR FILING DATE: 2002-04-01
; PRIOR FILING DATE: 2002-04-01
```

```
; PRIOR APPLICATION NUMBER: PCT/EP00/09657
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: A-1680/99
; PRIOR FILING DATE: 1999-10-01
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-11-082-595-20

Query Match      89.4%; Score 42; DB 20; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.6e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9
Db 1 DLMGYIPLV 9
|||||||

RESULT 31
US-08-854-825-1
; Sequence 1, Application US/08854825
; Publication No. US20020115061A1
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; APPLICANT: Cerny, Andreas
; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/854,825
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Silver, Donald J.
; REGISTRATION NUMBER: 37552
; REFERENCE/DOCKET NUMBER: 61230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-854-825-1

Query Match      89.4%; Score 42; DB 8; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9
Db 1 DLMGYIPLV 9
|||||||
```



```
Db      2 DLMGYIPLV 10

RESULT 32
US-10-777-053-109
; Sequence 109, Application US/10777053
; Publication No. US20040132089A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MAN/K.022C1
; CURRENT APPLICATION NUMBER: US/10/777,053
; CURRENT FILING DATE: 2004-02-10
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-10-777-053-109

Query Match      89.4%; Score 42; DB 16; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      2 DLMGYIPLV 10

RESULT 33
US-10-837-217-109
; Sequence 109, Application US/10837217
; Publication No. US20040203051A1
; GENERAL INFORMATION:
; APPLICANT: Simard, John J. L.
; APPLICANT: Diamond, David C.
; APPLICANT: Qiu, Zhiyong
; APPLICANT: Lei, Xiang-Dong
; TITLE OF INVENTION: EXPRESSION VECTORS ENCODING EPITOPES OF
; TARGET-ASSOCIATED ANTIGENS AND METHODS FOR THEIR DESIGN
; FILE REFERENCE: MAN/K.022C2
; CURRENT APPLICATION NUMBER: US/10/837,217
; CURRENT FILING DATE: 2004-04-30
; PRIOR APPLICATION NUMBER: 10/292,413
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: 60/336,968
; PRIOR FILING DATE: 2001-11-07
; NUMBER OF SEQ ID NOS: 979
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Hepatitis C Virus
US-10-837-217-109

Query Match      89.4%; Score 42; DB 16; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      2 DLMGYIPLV 10

RESULT 34
US-10-931-566-26
; Sequence 26, Application US/10931566
; Publication No. US20050106137A1
; GENERAL INFORMATION:
; APPLICANT: Apton Corporation
; APPLICANT: Grimes, Stephen
; TITLE OF INVENTION: Chimeric Peptide Immunogens
; FILE REFERENCE: 1102865-0047CIP
; CURRENT APPLICATION NUMBER: US/10/931,566
; CURRENT FILING DATE: 2004-08-31
; PRIOR APPLICATION NUMBER: 09/848,834
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: 60/202,328
; PRIOR FILING DATE: 2000-05-05
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-931-566-26

Query Match      89.4%; Score 42; DB 17; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.23;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      2 DLMGYIPLV 10

RESULT 35
US-10-103-395-172
; Sequence 172, Application US/10103395
; Publication No. US20020160019A1
; GENERAL INFORMATION:
; APPLICANT: EPIMUNE, Inc.
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; TITLE OF INVENTION: IDENTIFICATION OF BROADLY REACTIVE DR
; TITLE OF INVENTION: RESTRICTED EPITOPES
; FILE REFERENCE: 39963-20016.01
; CURRENT APPLICATION NUMBER: US/10/103,395
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: US 09/009,953
; PRIOR FILING DATE: 1998-01-21
; PRIOR APPLICATION NUMBER: PCT/US98/01373
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: US 60/036,713
; PRIOR FILING DATE: 1997-01-23
; PRIOR APPLICATION NUMBER: US 60/037,432
; PRIOR FILING DATE: 1997-02-07
; NUMBER OF SEQ ID NOS: 274
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-395-172

Query Match      89.4%; Score 42; DB 13; Length 15;
Best Local Similarity 88.9%; Pred. No. 0.35;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      2 DLMGYIPLV 10

RESULT 36
US-10-268-561-14
```

```

; ;
; FILE REFERENCE: J9963-Z0033.00
; CURRENT APPLICATION NUMBER: US/09/894,018
; ;
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; AFFILIANT: MARIENS, GEELC
; ;
; TITLE OF INVENTION: VIRUS
; CURRENT FILING DATE: 2001-06-27

```

Tue Aug 23 14:25:42 2005

```
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/10/651,165
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: US/08/974,690C
; PRIOR FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-10-651-165-10

Query Match      89.4%; Score 42; DB 15; Length 20;
Best Local Similarity 88.9%; Pred. No. 0.47;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
        |||||
Db      12 DLMGYIPLV 20
```

Search completed: August 23, 2005, 13:48:40  
Job time : 160 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2005, 13:23:38 : Search time 41 Seconds  
(without alignments)  
16.386 Million cell updates/sec

Title: US-10-770-117-1

Perfect score: 47

Sequence: 1 DLMGYIPAV 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 272444

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	100.0	9	4	US-09-763-260-1
2	42	89.4	9	1	US-08-214-650-54
3	42	89.4	9	4	US-08-197-484-51
4	42	89.4	9	4	US-08-197-484-130
5	42	89.4	9	4	US-09-311-784A-255
6	42	89.4	9	4	US-09-763-260-3
7	42	89.4	9	5	PCT-US95-02121-51
8	42	89.4	9	5	PCT-US95-02121-130
9	42	89.4	10	1	US-08-214-650-1
10	42	89.4	15	4	US-09-089-953-172
11	42	89.4	16	2	US-08-986-234-34
12	42	89.4	20	4	US-08-635-886C-10
13	42	89.4	20	4	US-08-974-690C-10
14	42	89.4	20	4	US-08-974-685-10
15	42	89.4	24	4	US-08-635-886C-216
16	42	89.4	24	4	US-08-974-690C-216
17	42	89.4	25	4	US-10-173-480-58
18	42	89.4	25	4	US-10-173-480-59
19	42	89.4	37	4	US-08-635-886C-281
20	42	89.4	37	4	US-08-974-690C-281
21	39	83.0	8	3	US-09-051-006-10
22	38	80.9	24	4	US-08-635-886C-285
23	38	80.9	24	4	US-08-974-690C-285
24	36	76.6	8	4	US-08-635-886C-69
25	36	76.6	8	4	US-08-974-690C-69
26	36	76.6	8	4	US-08-974-685-69
27	36	76.6	9	4	US-08-635-886C-163

28	36	76.6	9	4	US-08-974-690C-163	Sequence 163, App
29	36	76.6	9	4	US-08-974-685-163	Sequence 163, App
30	36	76.6	10	2	US-08-986-234-35	Sequence 35, App
31	36	76.6	20	4	US-08-635-886C-11	Sequence 11, Appl
32	36	76.6	20	4	US-08-974-690C-11	Sequence 11, Appl
33	36	76.6	20	4	US-08-974-685-11	Sequence 11, Appl
34	36	76.6	44	4	US-08-635-886C-67	Sequence 67, Appl
35	36	76.6	44	4	US-08-974-690C-67	Sequence 67, Appl
36	36	76.6	44	4	US-08-974-685-67	Sequence 67, Appl
37	35	74.5	23	1	US-08-189-331-179	Sequence 179, App
38	35	74.5	23	2	US-08-471-068-179	Sequence 179, App
39	33	70.2	28	4	US-09-091-097-45	Sequence 45, Appl
40	32	68.1	7	4	US-08-635-886C-70	Sequence 70, Appl
41	32	68.1	7	4	US-08-974-690C-70	Sequence 70, Appl
42	32	68.1	7	4	US-08-974-685-70	Sequence 70, Appl
43	31	66.0	20	1	US-07-678-974D-61	Sequence 61, Appl
44	31	66.0	20	2	US-08-945-168-67	Sequence 67, Appl
45	31	66.0	50	4	US-09-621-976-5926	Sequence 5926, Ap
46	29	61.7	24	4	US-08-492-723-3	Sequence 3, Appl1
47	28	59.6	9	4	US-09-239-043D-2277	Sequence 2277, Ap
48	28	59.6	10	3	US-09-385-442-20	Sequence 20, Appl
49	28	59.6	11	4	US-09-311-784A-284	Sequence 284, App
50	28	59.6	15	2	US-08-480-473B-15	Sequence 15, Appl
51	28	59.6	15	3	US-09-915-213-15	Sequence 15, Appl
52	28	59.6	15	3	US-09-235-217-15	Sequence 15, Appl
53	28	59.6	15	5	PCT-US96-10251-15	Sequence 15, Appl
54	28	59.6	20	1	US-07-678-974D-62	Sequence 62, Appl
55	28	59.6	20	2	US-08-945-168-68	Sequence 68, Appl
56	28	59.6	23	1	US-08-582-257-2	Sequence 2, Appl1
57	28	59.6	23	2	US-08-582-258-2	Sequence 2, Appl1
58	27	57.4	46	4	US-09-079-830-18	Sequence 18, Appl
59	27	57.4	50	3	US-08-938-830-5	Sequence 5, Appl1
60	27	57.4	50	3	US-09-020-223-5	Sequence 5, Appl1
61	26	55.3	15	4	US-09-009-953-176	Sequence 176, App
62	26	55.3	16	4	US-08-635-886C-61	Sequence 61, Appl
63	26	55.3	16	4	US-08-974-690C-61	Sequence 61, Appl
64	26	55.3	16	4	US-08-974-685-61	Sequence 61, Appl
65	26	55.3	16	4	US-08-974-685-62	Sequence 62, Appl
66	26	55.3	20	4	US-09-790-497A-446	Sequence 446, App
67	26	55.3	20	4	US-09-576-824A-446	Sequence 446, App
68	26	55.3	20	4	US-09-413-564C-35	Sequence 35, Appl
69	26	55.3	20	4	US-09-413-564C-36	Sequence 36, Appl
70	26	55.3	20	4	US-09-413-564C-37	Sequence 37, Appl
71	26	55.3	27	4	US-09-413-564C-34	Sequence 34, Appl
72	26	55.3	45	6	5177197-47	Patent No. 5177197
73	26	55.3	45	6	5177197-47	Patent No. 5177197
74	26	55.3	47	4	US-08-826-134-14	Sequence 14, Appl
75	26	55.3	47	4	US-08-826-134-25	Sequence 25, Appl
76	26	55.3	50	4	US-09-270-767-37069	Sequence 37069, A
77	26	55.3	50	4	US-09-270-767-52286	Sequence 52286, A
78	25	53.2	10	1	US-07-841-662-24	Sequence 24, Appl
79	25	53.2	10	1	US-08-209-797-24	Sequence 24, Appl
80	25	53.2	10	2	US-08-669-685-24	Sequence 24, Appl
81	25	53.2	10	3	US-09-103-486-24	Sequence 24, Appl
82	25	53.2	10	3	US-09-385-442-2	Sequence 2, Appl1
83	25	53.2	10	3	US-09-039-982A-40	Sequence 40, Appl
84	25	53.2	10	3	US-09-039-641-40	Sequence 40, Appl
85	25	53.2	10	3	US-09-039-762A-40	Sequence 40, Appl
86	25	53.2	10	3	US-09-042-492D-40	Sequence 40, Appl
87	25	53.2	10	4	US-08-913-612A-40	Sequence 40, Appl
88	25	53.2	10	4	US-10-266-463A-40	Sequence 40, Appl
89	25	53.2	10	5	PCT-US93-01557-24	Sequence 24, Appl
90	25	53.2	11	3	US-09-298-924-34	Sequence 34, Appl
91	25	53.2	13	2	US-08-210-421-1	Sequence 1, Appl1
92	25	53.2	13	3	US-08-704-727-1	Sequence 1, Appl1
93	25	53.2	13	3	US-09-412-420-1	Sequence 1, Appl1
94	25	53.2	13	5	PCT-US95-03311-1	Sequence 1, Appl1
95	25	53.2	20	1	US-08-399-646-9	Sequence 9, Appl1
96	25	53.2	20	1	US-08-607-321-9	Sequence 9, Appl1
97	25	53.2	20	2	US-08-961-240-9	Sequence 9, Appl1
98	25	53.2	20	2	US-08-605-501-9	Sequence 9, Appl1
99	25	53.2	25	1	US-08-433-854-10	Sequence 10, Appl
100	25	53.2	25	1	US-08-174-745A-10	Sequence 10, Appl

## ALIGNMENTS

RESULT 1  
US-09-763-260-1  
; Sequence 1, Application US/09763260  
; Patent No. 6685944  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States, as represented by the  
; APPLICANT: Secretary, Dept. of  
; APPLICANT: Health and Human Services  
; APPLICANT: Berzofsky, Jay A.  
; APPLICANT: Sarobe, Pablo  
; APPLICANT: Pendleton, C. David  
; APPLICANT: Feinstein, Stephen M.  
; APPLICANT: Arichi, Tatsumi  
; APPLICANT: Major, Marian E.  
; TITLE OF INVENTION: MODIFIED HCV PEPTIDE VACCINES  
; FILE REFERENCE: 14014.0347/P  
; CURRENT APPLICATION NUMBER: US/09/763.260  
; CURRENT FILING DATE: 2001-10-19  
; PRIOR APPLICATION NUMBER: 60/097,446  
; PRIOR FILING DATE: 1998-08-21  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:/No. 6685944e =  
; OTHER INFORMATION: synthetic construct  
US-09-763-260-1

Query Match 100.0%; Score 47; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
Db 1 DLMGYIPAV 9

RESULT 2  
US-08-214-650-54  
; Sequence 54, Application US/08214650  
; Patent No. 5703995  
; GENERAL INFORMATION:  
; APPLICANT: Chisari, Francis V.  
; APPLICANT: Cerny, Andreas  
; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T  
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS C VIRUS  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Leydig, Voit & Mayer  
; STREET: Two Prudential Plaza, Suite 4900  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/214,650  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Silver, Donald J.

; REGISTRATION NUMBER: 37552  
; REFERENCE/DOCKET NUMBER: 61230  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 616-5600  
; TELEFAX: (312) 616-5700  
; TELEX: 25-3533  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-214-650-54  
Query Match 89.4%; Score 42; DB 1; Length 9;  
Best Local Similarity 88.9%; Pred. No. 4.1e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 DLMGYIPAV 9  
Db 1 DLMGYIPLV 9  
RESULT 3  
US-08-197-484-51  
; Sequence 51, Application US/08197484  
; Patent No. 6419931  
; GENERAL INFORMATION:  
; APPLICANT: VITIELLO, Maria A.  
; APPLICANT: CHESTNUT, Robert W.  
; APPLICANT: SETTE, Alessandro D.  
; APPLICANT: CELIS, Eteban  
; APPLICANT: GRAY, Howard  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; TITLE OF INVENTION: CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/197,484  
; FILING DATE: 16-FEB-1994  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/935,811  
; FILING DATE: 26-AUG-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/874,491  
; FILING DATE: 27-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/827,682  
; FILING DATE: 29-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/749,568  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Parmelee, Steven W.  
; REGISTRATION NUMBER: 31,990  
; REFERENCE/DOCKET NUMBER: 14137-26-4  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 467-9600  
; TELEFAX: (206) 623-6793  
; INFORMATION FOR SEQ ID NO: 51:

SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-197-484-51

Query Match 89.4%; Score 42; DB 4; Length 9;  
Best Local Similarity 88.9%; Pred. No. 4.1e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
Db 1 DLMGYIPLV 9

RESULT 4  
US-08-197-484-130  
; Sequence 130, Application US/08197484  
; Patent No. 6419331  
; GENERAL INFORMATION:  
; APPLICANT: VITIELLO, Maria A.  
; APPLICANT: CHESTNUT, Robert W.  
; APPLICANT: SETTE, Alessandro D.  
; APPLICANT: CELIS, Esteban  
; APPLICANT: GRAY, Howard  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING  
; TITLE OF INVENTION: CTL IMMUNITY  
; NUMBER OF SEQUENCES: 153  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Khourie and Crew  
; STREET: Steuart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/197,484  
FILING DATE: 16-FEB-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/935,811  
FILING DATE: 26-AUG-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/874,491  
FILING DATE: 27-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/827,682  
FILING DATE: 29-JAN-1992

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/749,568  
FILING DATE: 26-AUG-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.

REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 14137-26-4  
TELEPHONE: (206) 467-9600  
TELEFAX: (206) 623-6793

INFORMATION FOR SEQ ID NO: 130:

SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-197-484-130

Query Match 89.4%; Score 42; DB 4; Length 9;  
Best Local Similarity 88.9%; Pred. No. 4.1e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
Db 1 DLMGYIPLV 9

RESULT 5  
US-09-311-784A-255  
; Sequence 255, Application US/09311784A  
; Patent No. 6534482  
; GENERAL INFORMATION:  
; APPLICANT: Fikes, John D.  
; APPLICANT: Hermanson, Gary G.  
; APPLICANT: Sette, Alessandro  
; APPLICANT: Ishioka, Glenn Y.  
; APPLICANT: Livingston, Brian  
; APPLICANT: Chesnut, Robert W.  
; APPLICANT: Epimmune Inc.  
; TITLE OF INVENTION: Expression Vectors for Stimulating an  
; TITLE OF INVENTION: Immune Response and Methods of Using the Same  
; FILE REFERENCE: 39963-20022.01  
; CURRENT APPLICATION NUMBER: US/09/311,784A  
; CURRENT FILING DATE: 1999-05-13  
; PRIOR APPLICATION NUMBER: US 60/085,751  
; PRIOR FILING DATE: 1998-05-15  
; NUMBER OF SEQ ID NOS: 463  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 255  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HCV Core 132 (peptide 1013.1002)  
US-09-311-784A-255

Query Match 89.4%; Score 42; DB 4; Length 9;  
Best Local Similarity 88.9%; Pred. No. 4.1e+05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
Db 1 DLMGYIPLV 9

RESULT 6  
US-09-763-260-3  
; Sequence 3, Application US/09763260  
; Patent No. 6685944  
; GENERAL INFORMATION:  
; APPLICANT: The Government of the United States, as represented by the  
; APPLICANT: Secretary, Dept. of  
; APPLICANT: Health and Human Services  
; APPLICANT: Berzofsky, Jay A.  
; APPLICANT: Sarobe, Pablo  
; APPLICANT: Pendleton, C. David  
; APPLICANT: Feinstein, Stephen M.  
; APPLICANT: Arichi, Tatsuami  
; APPLICANT: Major, Marian E.  
; TITLE OF INVENTION: MODIFIED HCV PEPTIDE VACCINES  
; FILE REFERENCE: 14014.0347/P  
; CURRENT APPLICATION NUMBER: US/09/763,260  
; CURRENT FILING DATE: 2001-10-19  
; PRIOR APPLICATION NUMBER: 60/097,446  
; PRIOR FILING DATE: 1998-08-21  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 9  
; TYPE: PRT

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence./No. 6685944e =
US-09-763-260-3
Query Match      89.4%; Score 42; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9
   |||||
Db 1 DLMGYIPLV 9

RESULT 7
PCT-US95-02121-51
; Sequence 51, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US95-02121-51

Query Match      89.4%; Score 42; DB 5; Length 9;
Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9
   |||||
Db 1 DLMGYIPLV 9

RESULT 8
; Sequence 51, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US95-02121-51

Query Match      89.4%; Score 42; DB 5; Length 9;
Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9
   |||||
Db 1 DLMGYIPLV 9

RESULT 9
US-08-214-650-1
; Sequence 1, Application US/08214650
; Patent No. 5709995
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; APPLICANT: Cerny, Andreas
; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voigt & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601

```

```

PCT-US95-02121-130
; Sequence 130, Application PC/TUS9502121
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR ELICITING
; TITLE OF INVENTION: CTL IMMUNITY
; NUMBER OF SEQUENCES: 153
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02121
; FILING DATE: 16-FEB-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/197,484
; FILING DATE: 16-FEB-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/935,811
; FILING DATE: 26-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/874,491
; FILING DATE: 27-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/827,682
; FILING DATE: 29-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/749,568
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 14137-26-4PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 130:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US95-02121-130

Query Match      89.4%; Score 42; DB 5; Length 9;
Best Local Similarity 88.9%; Pred. No. 4.1e+05;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9
   |||||
Db 1 DLMGYIPLV 9

RESULT 9
US-08-214-650-1
; Sequence 1, Application US/08214650
; Patent No. 5709995
; GENERAL INFORMATION:
; APPLICANT: Chisari, Francis V.
; APPLICANT: Cerny, Andreas
; TITLE OF INVENTION: PEPTIDES FOR INDUCING CYTOTOXIC T
; TITLE OF INVENTION: LYMPHOCYTE RESPONSES TO HEPATITIS C VIRUS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voigt & Mayer
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601

```



```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/214,650
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Silver, Donald J.
; REGISTRATION NUMBER: 37552
; REFERENCE/DOCKET NUMBER: 61230
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-214-650-1

Query Match      89.4%; Score 42; DB 1; Length 10;
Best Local Similarity 88.9%; Pred. No. 0.023;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      2 DLMGYIPLV 10

RESULT 10
US-09-009-953-172
; Sequence 172, Application US/09009953
; Patent No. 6413517
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; TITLE OF INVENTION: Identification of Broadly
;             Reactive DR Restricted Epitopes
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,953
; FILING DATE: 21-Jan-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,713
; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 60/037,432
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-011520US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; TELEX: <Unknown>
```

```
; INFORMATION FOR SEQ ID NO: 172:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 172:
US-09-009-953-172

Query Match      89.4%; Score 42; DB 4; Length 15;
Best Local Similarity 88.9%; Pred. No. 0.036;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      2 DLMGYIPLV 10

RESULT 11
US-08-986-234-34
; Sequence 34, Application US/08986234
; Patent No. 5981706
; GENERAL INFORMATION:
; APPLICANT: Wallen, et al.
; TITLE OF INVENTION: Methods for Synthesizing Heat Shock Protein Complexes
; FILE REFERENCE: UNME-0008-1
; CURRENT APPLICATION NUMBER: US/08/986,234
; CURRENT FILING DATE: 1997-12-05
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 34
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; US-08-986-234-34

Query Match      89.4%; Score 42; DB 2; Length 16;
Best Local Similarity 88.9%; Pred. No. 0.039;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      4 DLMGYIPLV 12

RESULT 12
US-08-635-886C-10
; Sequence 10, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635,886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
; US-08-635-886C-10

Query Match      89.4%; Score 42; DB 4; Length 20;
Best Local Similarity 88.9%; Pred. No. 0.05;
```

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | | |  
Db 12 DLMGYIPLV 20

RESULT 13  
US-08-974-690C-10  
; Sequence 10, Application US/08974690C  
; Patent No. 6613333  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; TITLE OF INVENTION: VIRUS  
; FILE REFERENCE: 2551-94  
; CURRENT APPLICATION NUMBER: US/08/974,690C  
; CURRENT FILING DATE: 1997-11-19  
; PRIOR FILING DATE: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-08-974-690C-10

Query Match 89.4%; Score 42; DB 4; Length 20;  
Best Local Similarity 88.9%; Pred. No. 0.05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | | |  
Db 12 DLMGYIPLV 20

RESULT 14  
US-08-974-685-10  
; Sequence 10, Application US/08974685  
; Patent No. 6689368  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, GEERT  
; APPLICANT: DELEYS, ROBERT  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF  
HEPATITIS C VIRUS STRUCTURAL PROTEINS, PROPHYLACTIC AND  
THERAPEUTIC COMPOSITIONS CONTAINING THE SAME

NUMBER OF SEQUENCES: 181  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHVE P.C.  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,685  
FILING DATE: 19-NOV-97  
CLASSIFICATION: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BYRNE, THOMAS E.  
REGISTRATION NUMBER: 32,205  
REFERENCE/DOCKET NUMBER: 1487-11

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-08-974-685-10

Query Match 89.4%; Score 42; DB 4; Length 20;  
Best Local Similarity 88.9%; Pred. No. 0.05;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | | |  
Db 12 DLMGYIPLV 20

RESULT 15  
US-08-635-886C-216  
; Sequence 216, Application US/08635886C  
; Patent No. 6555114  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; TITLE OF INVENTION: VIRUS  
; FILE REFERENCE: 2752-18  
; CURRENT APPLICATION NUMBER: US/08/635,886C  
; CURRENT FILING DATE: 1996-04-25  
; PRIOR FILING DATE: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 216  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-08-635-886C-216

Query Match 89.4%; Score 42; DB 4; Length 24;  
Best Local Similarity 88.9%; Pred. No. 0.062;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | | |  
Db 6 DLMGYIPLV 14

RESULT 16  
US-08-974-690C-216  
; Sequence 216, Application US/08974690C  
; Patent No. 6613333  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; TITLE OF INVENTION: VIRUS  
; FILE REFERENCE: 2551-94  
; CURRENT APPLICATION NUMBER: US/08/974,690C  
; CURRENT FILING DATE: 1997-11-19  
; PRIOR FILING DATE: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04

Tue Aug 23 14:25:42 2005

; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 216  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-08-974-690C-216

Query Match 89.4%; Score 42; DB 4; Length 24;  
Best Local Similarity 88.9%; Pred. No. 0.062;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | | |  
Db 6 DLMGYIPLV 14

RESULT 17  
US-10-173-480-58 Application US/10173480

; Sequence 58, Application US/10173480  
; Patent No. 6727092  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Shah, Dinesh O.  
; APPLICANT: Dawson, George A.  
; APPLICANT: Muerhoff, A. Scott  
; APPLICANT: Jiang, Lily  
; APPLICANT: Gutierrez, Robin A.  
; APPLICANT: Leary, Thomas P.  
; APPLICANT: Desai, Suresh  
; APPLICANT: Stewart, James L.  
; TITLE OF INVENTION: METHODS FOR THE SIMULTANEOUS DETECTION  
; OF HCV ANTIGENS AND HCV ANTIBODIES  
; FILE REFERENCE: 6821.US.P1  
; CURRENT APPLICATION NUMBER: US/10/173,480  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 09/891,983  
; PRIOR FILING DATE: 2001-06-26  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 58  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HCV-Core derived peptides  
US-10-173-480-58

Query Match 89.4%; Score 42; DB 4; Length 25;  
Best Local Similarity 88.9%; Pred. No. 0.065;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | | |  
Db 17 DLMGYIPLV 25

RESULT 18  
US-10-173-480-59 Application US/10173480

; Sequence 59, Application US/10173480  
; Patent No. 6727092  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Shah, Dinesh O.  
; APPLICANT: Dawson, George A.  
; APPLICANT: Muerhoff, A. Scott  
; APPLICANT: Jiang, Lily  
; APPLICANT: Gutierrez, Robin A.  
; APPLICANT: Leary, Thomas P.  
; APPLICANT: Desai, Suresh  
; APPLICANT: Stewart, James L.  
; TITLE OF INVENTION: METHODS FOR THE SIMULTANEOUS DETECTION  
; OF HCV ANTIGENS AND HCV ANTIBODIES

; FILE REFERENCE: 6821.US.P1  
; CURRENT APPLICATION NUMBER: US/10/173,480  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 09/891,983  
; PRIOR FILING DATE: 2001-06-26  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 59  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: HCV-Core derived peptides  
US-10-173-480-59

Query Match 89.4%; Score 42; DB 4; Length 25;  
Best Local Similarity 88.9%; Pred. No. 0.065;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | | |  
Db 7 DLMGYIPLV 15

RESULT 19  
US-08-635-886C-281 Application US/08635886C  
; Sequence 281, Application US/08635886C  
; Patent No. 6555114  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; VIRUS  
; FILE REFERENCE: 2752-18  
; CURRENT APPLICATION NUMBER: US/08/635,886C  
; CURRENT FILING DATE: 1996-04-25  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 281  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Xaa is any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (14)..(14)  
; OTHER INFORMATION: Xaa is any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (28)..(29)  
; OTHER INFORMATION: Xaa is any amino acid  
US-08-635-886C-281

Query Match 89.4%; Score 42; DB 4; Length 37;  
Best Local Similarity 88.9%; Pred. No. 0.1;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | | |  
Db 19 DLMGYIPLV 27

RESULT 20  
US-08-974-690C-281 Application US/08974690C  
; Sequence 281, Application US/08974690C

Patent No. 6613333  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; TITLE OF INVENTION: VIRUS  
; FILE REFERENCE: 2551-94  
; CURRENT APPLICATION NUMBER: US/08/974,690C  
; CURRENT FILING DATE: 1997-11-19  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 281  
; LENGTH: 37  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Xaa is any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (14)..(14)  
; OTHER INFORMATION: Xaa is any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (28)..(29)  
; OTHER INFORMATION: Xaa is any amino acid  
US-08-974-690C-281

Query Match 89.4%; Score 42; DB 4; Length 37;  
Best Local Similarity 89.9%; Pred. No. 0.1;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | | |  
Db 19 DLMGYIPLV 27

RESULT 21  
US-09-051-006-10  
; Sequence 10, Application US/09051006  
; Patent No. 6380359  
; GENERAL INFORMATION:  
; APPLICANT: Mogam Biotechnology Research Institute  
; APPLICANT: Kim, Tae-Young  
; APPLICANT: Lee, Ki-Young  
; APPLICANT: Chang, Jin-Soo  
; APPLICANT: Cho, Sung-Yoo  
; APPLICANT: Hwang, Yu-Kyeong  
; APPLICANT: Choi, Myeong  
; APPLICANT: Cheong, Hong-Seok  
; TITLE OF INVENTION: Liposomes Comprising Peptide Antigens  
; TITLE OF INVENTION: Derived from X Protein of Hepatitis B virus  
; FILE REFERENCE: 0136/OE154  
; CURRENT APPLICATION NUMBER: US/09/051,006  
; CURRENT FILING DATE: 1998-03-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 10  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: HCV  
US-09-051-006-10

Query Match 83.0%; Score 39; DB 3; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DLMGYIP 7  
| | | | | | | |  
Db 1 DLMGYIP 7

RESULT 22  
US-08-635-886C-285  
; Sequence 285, Application US/08635886C  
; Patent No. 6555114  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; TITLE OF INVENTION: VIRUS  
; FILE REFERENCE: 2752-18  
; CURRENT APPLICATION NUMBER: US/08/635,886C  
; CURRENT FILING DATE: 1996-04-25  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 285  
; LENGTH: 24  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (3)..(4)  
; OTHER INFORMATION: Xaa is any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (11)..(11)  
; OTHER INFORMATION: Xaa is any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (21)..(21)  
; OTHER INFORMATION: Xaa is any amino acid  
US-08-635-886C-285

Query Match 80.9%; Score 38; DB 4; Length 24;  
Best Local Similarity 77.8%; Pred. No. 0.41;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | | |  
Db 6 DLMGYXPVV 14

RESULT 23  
US-08-974-690C-285  
; Sequence 285, Application US/08974690C  
; Patent No. 6613333  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; TITLE OF INVENTION: VIRUS  
; FILE REFERENCE: 2551-94  
; CURRENT APPLICATION NUMBER: US/08/974,690C  
; CURRENT FILING DATE: 1997-11-19  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 285  
; LENGTH: 24  
; TYPE: PRT

; ORGANISM: hepatitis C virus  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (3)..(4)  
; OTHER INFORMATION: Xaa is any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (11)..(11)  
; OTHER INFORMATION: Xaa is any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (21)..(21)  
; OTHER INFORMATION: Xaa is any amino acid  
US-08-974-690C-285

Query Match 80.9%; Score 38; DB 4; Length 24;  
Best Local Similarity 77.8%; Pred. No. 0.41;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LMGVIVAV 9  
| | | | | | |  
Db 6 DLMGYXPV 14

RESULT 24  
US-08-635-886C-69  
; Sequence 69, Application US/08635886C  
; Patent No. 6555114  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; FILE REFERENCE: 2752-18  
; CURRENT APPLICATION NUMBER: US/08/635,886C  
; PRIORITY FILING DATE: 1996-04-25  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 69  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-08-635-886C-69

Query Match 76.6%; Score 36; DB 4; Length 8;  
Best Local Similarity 87.5%; Pred. No. 4.1e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGVIVAV 9  
| | | | | | |  
Db 1 LMGYIPLV 8

RESULT 25  
US-08-974-690C-69  
; Sequence 69, Application US/08974690C  
; Patent No. 6613333  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; FILE REFERENCE: 2551-94  
; CURRENT APPLICATION NUMBER: US/08/974,690C  
; CURRENT FILING DATE: 1997-11-19  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28

; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 69  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-08-974-690C-69

Query Match 76.6%; Score 36; DB 4; Length 8;  
Best Local Similarity 87.5%; Pred. No. 4.1e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGVIVAV 9  
| | | | | | |  
Db 1 LMGYIPLV 8

RESULT 26  
US-08-974-685-69  
; Sequence 69, Application US/08974685  
; Patent No. 668368  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, GEERT  
; APPLICANT: DELEYS, ROBERT  
; APPLICANT: MAERTENS, GEERT  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF  
HEPATITIS C VIRUS STRUCTURAL PROTEINS, PROPHYLACTIC AND  
THERAPEUTIC COMPOSITIONS CONTAINING THE SAME  
; NUMBER OF SEQUENCES: 181  
; CORRESPONDENCE ADDRESS:  
ADDRESSEE: NIXON & VANDERHYE P.C.  
STREET: 1100 NORTH GLEBE ROAD  
CITY: ARLINGTON  
STATE: VIRGINIA  
COUNTRY: U.S.A.  
ZIP: 22201-4714  
; COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
; CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,685  
FILING DATE: 19-NO. 668368-1997  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: BYRNE, THOMAS E.  
REGISTRATION NUMBER: 32,205  
REFERENCE/DOCKET NUMBER: 1487-11  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 816-4000  
TELEFAX: (703) 816-4100  
; INFORMATION FOR SEQ ID NO: 69:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 69:

Query Match 76.6%; Score 36; DB 4; Length 8;  
Best Local Similarity 87.5%; Pred. No. 4.1e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGVIVAV 9  
| | | | | | |  
Db 1 LMGYIPLV 8

```
RESULT 27
US-08-635-886C-163
; Sequence 163, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2752-18
; CURRENT FILING DATE: 1996-04-25
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 163
; LENGTH: 9
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-163

Query Match      76.6%; Score 36; DB 4; Length 9;
Best Local Similarity 87.5%; Pred. No. 4.1e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 LMGYIPAV 9
      |||||
Db      2 LMGYIPLV 9

RESULT 28
US-08-974-690C-163
; Sequence 163, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1993-11-04
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 163
; LENGTH: 9
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-163

Query Match      76.6%; Score 36; DB 4; Length 9;
Best Local Similarity 87.5%; Pred. No. 4.1e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 LMGYIPAV 9
      |||||
Db      2 LMGYIPLV 9

RESULT 29
US-08-974-685-163
; Sequence 163, Application US/08974685
; Patent No. 6689368
; GENERAL INFORMATION:
```

```
APPLICANT: LEROUX-ROELS, GEERT
DELEYS, ROBERT
MAERTENS, GEERT
TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF
HEPATITIS C VIRUS STRUCTURAL PROTEINS, PROPHYLACTIC AND
THERAPEUTIC COMPOSITIONS CONTAINING THE SAME
NUMBER OF SEQUENCES: 181
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/974,685
FILING DATE: 19-NO. 6689368-1997
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: BYRNE, THOMAS E.
REGISTRATION NUMBER: 32,205
REFERENCE/DOCKET NUMBER: 1487-11
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 163:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 163:
US-08-974-685-163

Query Match      76.6%; Score 36; DB 4; Length 9;
Best Local Similarity 87.5%; Pred. No. 4.1e+05;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 LMGYIPAV 9
      |||||
Db      2 LMGYIPLV 9

RESULT 30
US-08-986-234-35
; Sequence 35, Application US/08986234
; Patent No. 5981706
; GENERAL INFORMATION:
; APPLICANT: Wallen, et al.
; TITLE OF INVENTION: Methods for Synthesizing Heat Shock Protein Complexes
; FILE REFERENCE: UNME-0008-1
; CURRENT APPLICATION NUMBER: US/08/986,234
; CURRENT FILING DATE: 1997-12-05
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-08-986-234-35

Query Match      76.6%; Score 36; DB 2; Length 10;
Best Local Similarity 87.5%; Pred. No. 0.4;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2 LMGYIPAV 9
      |||||
```

```
Db      1 LMGVIPLV 8

RESULT 31
US-08-635-886C-11
; Sequence 11, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635.886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-11

Query Match      76.6%; Score 36; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.87;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy      2 LMGYIPAV 9
       |||||
Db      1 LMGVIPLV 8

RESULT 32
US-08-974-690C-11
; Sequence 11, Application US/08974690C
; Patent No. 6613333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974.690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11
; LENGTH: 20
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-11

Query Match      76.6%; Score 36; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.87;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy      2 LMGYIPAV 9
       |||||
Db      1 LMGVIPLV 8

RESULT 33
US-08-974-685-11
; Sequence 11, Application US/08974685
; Patent No. 6689368
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, GEERT
; APPLICANT: DELEYS, ROBERT
; APPLICANT: MAERTENS, GEERT
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF
; HEPATITIS C VIRUS STRUCTURAL PROTEINS, PROPHYLACTIC AND
; THERAPEUTIC COMPOSITIONS CONTAINING THE SAME
; NUMBER OF SEQUENCES: 181
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974.685
; FILING DATE: 19-NO. 6689368-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-08-974-685-11

Query Match      76.6%; Score 36; DB 4; Length 20;
Best Local Similarity 87.5%; Pred. No. 0.87;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

Qy      2 LMGYIPAV 9
       |||||
Db      1 LMGVIPLV 8

RESULT 34
US-08-635-886C-67
; Sequence 67, Application US/08635886C
; Patent No. 6555114
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; APPLICANT: MAERTENS, Geert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; TITLE OF INVENTION: VIRUS
; FILE REFERENCE: 2752-18
; CURRENT APPLICATION NUMBER: US/08/635.886C
; CURRENT FILING DATE: 1996-04-25
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 67
; LENGTH: 44
```

```
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-635-886C-67

Query Match      76.6%; Score 36; DB 4; Length 44;
Best Local Similarity 87.5%; Pred. No. 2.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGYIPAV 9
Db 1 LMGYIPLV 8

RESULT 35
US-08-974-690C-67
; Sequence 67, Application US/08974690C
; Patent No. 661333
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, Geert
; APPLICANT: DELEYS, Robert
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C
; FILE REFERENCE: 2551-94
; CURRENT APPLICATION NUMBER: US/08/974,690C
; CURRENT FILING DATE: 1997-11-19
; PRIOR APPLICATION NUMBER: PCT/EP94/03555
; PRIOR FILING DATE: 1994-10-28
; PRIOR APPLICATION NUMBER: EP 93402718.6
; PRIOR FILING DATE: 1993-11-04
; NUMBER OF SEQ ID NOS: 286
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 44
; TYPE: PRT
; ORGANISM: hepatitis C virus
US-08-974-690C-67

Query Match      76.6%; Score 36; DB 4; Length 44;
Best Local Similarity 87.5%; Pred. No. 2.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGYIPAV 9
Db 1 LMGYIPLV 8

RESULT 36
US-08-974-685-67
; Sequence 67, Application US/08974685
; Patent No. 6689368
; GENERAL INFORMATION:
; APPLICANT: LEROUX-ROELS, GEERT
; APPLICANT: DELEYS, ROBERT
; APPLICANT: MAERTENS, GEERT
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF
; HEPATITIS C VIRUS STRUCTURAL PROTEINS, PROPHYLACTIC AND
; THERAPEUTIC COMPOSITIONS CONTAINING THE SAME
; NUMBER OF SEQUENCES: 181
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/974,685
```

```
; FILING DATE: 19-No. 6689368-1997
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BYRNE, THOMAS E.
; REGISTRATION NUMBER: 32,205
; REFERENCE/DOCKET NUMBER: 1487-11
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 67:
US-08-974-685-67

Query Match      76.6%; Score 36; DB 4; Length 44;
Best Local Similarity 87.5%; Pred. No. 2.1;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 LMGYIPAV 9
Db 1 LMGYIPLV 8

RESULT 37
US-08-189-331-179
; Sequence 179, Application US/08189331
; Patent No. 5747334
; GENERAL INFORMATION:
; APPLICANT: Kay, B. K.
; APPLICANT: Fowlkes, D. M.
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents
; NUMBER OF SEQUENCES: 186
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/189,331
; FILING DATE: Concurrently herewith
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-155
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 179:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-189-331-179

Query Match      74.5%; Score 35; DB 1; Length 23;
Best Local Similarity 75.0%; Pred. No. 1.6;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```



Qy 1 DLMGVIPA 8  
Db 16 DLSGYLPA 23

RESULT 38  
US-08-471-068-179  
; Sequence 179, Application US/08471068  
; Patent No. 5948635  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 186  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471.068  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/189,331  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-155  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 179:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-471-068-179

Query Match 74.5%; Score 35; DB 2; Length 23;  
Best Local Similarity 75.0%; Pred. No. 1.6;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGVIPA 8  
Db 16 DLSGYLPA 23

RESULT 39  
US-09-091-097-45  
; Sequence 45, Application US/09091097  
; Patent No. 6432407  
; GENERAL INFORMATION:  
; APPLICANT: TAKESAKO, KAZUTOH  
; APPLICANT: OKADO, TAKASHI  
; APPLICANT: YAGIHARA, TOMOKO  
; APPLICANT: KURODA, MASANOBU  
; APPLICANT: ONISHI, YOSHIMI  
; APPLICANT: KATO, IKUNOSHIN  
; APPLICANT: AKIYAMA, KAZUO  
; APPLICANT: YASUEDA, HIROSHI

; APPLICANT: YAMAGUCHI, HIDEYO  
; TITLE OF INVENTION: ANTIGENIC PROTEIN ORIGINATING IN  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP  
; STREET: PO BOX 747  
; CITY: FALLS CHURCH  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/091,097  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: WEINER, MARC S.  
; REGISTRATION NUMBER: 32,181  
; REFERENCE/DOCKET NUMBER: 1422-0346P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-205-8000  
; TELEFAX: 703-205-8050  
; INFORMATION FOR SEQ ID NO: 45:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 28 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-091-097-45

Query Match 70.2%; Score 33; DB 4; Length 28;  
Best Local Similarity 100.0%; Pred. No. 5.3;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LMGVIP 7  
Db 17 LMGVIP 22

RESULT 40  
US-08-635-886C-70  
; Sequence 70, Application US/08635886C  
; Patent No. 6555114  
; GENERAL INFORMATION:  
; APPLICANT: LEROUX-ROELS, Geert  
; APPLICANT: DELEYS, Robert  
; APPLICANT: MAERTENS, Geert  
; TITLE OF INVENTION: IMMUNODOMINANT HUMAN T CELL EPITOPES OF HEPATITIS C  
; FILE REFERENCE: 2752-18  
; CURRENT APPLICATION NUMBER: US/08/635,886C  
; CURRENT FILING DATE: 1996-04-25  
; PRIOR APPLICATION NUMBER: PCT/EP94/03555  
; PRIOR FILING DATE: 1994-10-28  
; PRIOR APPLICATION NUMBER: EP 93402718.6  
; PRIOR FILING DATE: 1993-11-04  
; NUMBER OF SEQ ID NOS: 286  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 70  
; LENGTH: 7  
; TYPE: PRT  
; ORGANISM: hepatitis C virus  
US-08-635-886C-70

Query Match 68.1%; Score 32; DB 4; Length 7;  
Best Local Similarity 85.7%; Pred. No. 4.1e+05;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 MGYIPAV 9  
| | | | |  
Db 1 MGYIPLV 7

Search completed: August 23, 2005, 13:35:47  
Job time : 42 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 23, 2005, 13:17:43 ; Search time 161 Seconds  
(without alignments)

21.620 Million cell updates/sec

Title: US-10-770-117-1

Perfect score: 47

Sequence: 1 DLNGYIPAV 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 938430

Minimum DB seq length: 0

Maximum DB seq length: 50

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : A Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	47	100.0	9	3	AAV82996
2	47	100.0	9	4	AB82052
3	43	91.5	15	8	ADL25899
4	43	91.5	15	8	ADL25904
5	43	91.5	15	8	ADL25905
6	43	91.5	15	8	ADL25900
7	43	91.5	15	8	ADL25901
8	43	91.5	15	8	ADL25903
9	43	91.5	15	8	ADL25902
10	43	91.5	36	8	ADL26466
11	42	89.4	9	2	AAV73111
12	42	89.4	9	2	AAV59125
13	42	89.4	9	2	AAV78946
14	42	89.4	9	2	AAV87652
15	42	89.4	9	2	AAV91054
16	42	89.4	9	2	AAV39446
17	42	89.4	9	2	AAV54635
18	42	89.4	9	2	AAV10219
19	42	89.4	9	2	AAV10441
20	42	89.4	9	3	AAV73106
21	42	89.4	9	3	AAV82998
22	42	89.4	9	4	AAJ03866
23	42	89.4	9	4	AAJ01746
24	42	89.4	9	4	AAJ03964
25	42	89.4	9	4	AAJ00011
26	42	89.4	9	4	AAJ00229
27	42	89.4	9	4	AAJ03787
28	42	89.4	9	4	AAJ04086
29	42	89.4	9	5	ABG79901
30	42	89.4	9	5	ABG80124
31	42	89.4	9	6	ABP70576
32	42	89.4	9	7	ADA49528
33	42	89.4	9	8	ADP75563
34	42	89.4	9	8	ADJ79563
35	42	89.4	9	8	ADK69063
36	42	89.4	9	8	ADK68837
37	42	89.4	9	8	ADJ84718
38	42	89.4	9	8	ADO24206
39	42	89.4	9	8	ADO08619
40	42	89.4	9	8	ADQ10635
41	42	89.4	9	8	ADQ10634
42	42	89.4	9	8	ADQ10861
43	42	89.4	9	8	ADR11676
44	42	89.4	9	8	ADR69426
45	42	89.4	10	2	AAW84570
46	42	89.4	10	2	AAW39445
47	42	89.4	10	2	AAV10135
48	42	89.4	10	5	ABG79818
49	42	89.4	10	8	ADK68746
50	42	89.4	10	8	ADQ10544
51	42	89.4	11	4	AAJ01550
52	42	89.4	11	4	AAJ00230
53	42	89.4	11	4	AAJ01947
54	42	89.4	11	4	AAJ02126
55	42	89.4	11	4	AAJ00276
56	42	89.4	15	2	AAW85270
57	42	89.4	15	4	AAJ03485
58	42	89.4	15	4	AAJ03147
59	42	89.4	15	4	AAJ03178
60	42	89.4	15	4	AAJ03985
61	42	89.4	15	7	ABR55585
62	42	89.4	15	7	ADH19101
63	42	89.4	15	8	ADI34998
64	42	89.4	15	8	ADI34999
65	42	89.4	15	8	ADL25886
66	42	89.4	15	8	ADL25882
67	42	89.4	15	8	ADL25885
68	42	89.4	15	8	ADL25883
69	42	89.4	15	8	ADL25887
70	42	89.4	15	8	ADL25888
71	42	89.4	15	8	ADL25884
72	42	89.4	15	8	ADO23855
73	42	89.4	16	2	AAV87651
74	42	89.4	16	2	AAV25210
75	42	89.4	16	6	ABP70575
76	42	89.4	20	2	AAV56611
77	42	89.4	20	2	AAV56610
78	42	89.4	20	2	AAV84500
79	42	89.4	20	4	AAV82051
80	42	89.4	20	4	AAV64436
81	42	89.4	20	7	ADA49671
82	42	89.4	20	8	ADO23856
83	42	89.4	21	4	AAV64435
84	42	89.4	21	7	ADA49670
85	42	89.4	22	5	ABW04527
86	42	89.4	23	8	ADI34954
87	42	89.4	25	6	ABJ37519
88	42	89.4	25	6	ABJ37520
89	42	89.4	27	8	ADJ84719
90	42	89.4	30	5	AAU84606
91	42	89.4	36	8	ADL26465
92	42	89.4	36	8	ADL26468
93	39	83.0	8	4	AAJ01618
94	39	83.0	8	4	AAJ00864
95	39	83.0	8	4	AAJ00228
96	39	83.0	9	8	ADN64405
97	39	83.0	10	4	AAJ01478
98	39	83.0	10	4	AAJ01478

AAJ00229 Hepatitis  
AAJ03787 Hepatitis  
AAJ04086 Hepatitis  
ABG79901 MHC class  
ABG80124 MHC class  
ABP70576 Hepatitis  
ADA49528 Multi-epi  
ADE97587 Immunogen  
ADK69063 CTL epitope  
ADK68837 Epitope 1  
ADJ84718 Hepatitis  
ADO24206 HCV epitope  
ADO08619 Hepatitis  
ADQ10635 Hepatitis  
ADQ10634 Human her  
ADQ10861 Hepatitis  
ADR11676 HLA-A2.1  
ADR69426 Novel hyb  
ADR84570 Cytotoxic  
AAW39445 Human HCV  
AAV10135 T cell ep  
ABG79818 MHC class  
ADK68746 Epitope 1  
ADQ10544 Hepatitis  
AAJ01550 Hepatitis  
AAJ00230 Hepatitis  
AAJ01947 Hepatitis  
AAJ02126 Hepatitis  
AAJ00276 Hepatitis  
AAW85270 Helper T-  
AAJ03485 Hepatitis  
AAJ03147 Hepatitis  
AAJ03178 Hepatitis  
AAJ03985 Hepatitis  
ABR55585 Amino aci  
ADH19101 HCV core  
ADI34998 Hepatitis  
ADI34999 Hepatitis  
ADL25886 Synthetic  
ADL25882 Synthetic  
ADL25885 Synthetic  
ADL25883 Synthetic  
ADL25887 Synthetic  
ADL25888 Synthetic  
ADL25884 Synthetic  
ADO23855 HBV pepti  
AAV87651 Hepatitis  
AAV25210 HCV viral  
ABP70575 Hepatitis  
AAV56611 HCV pepti  
AAV56610 HCV pepti  
AAV84500 Hepatitis  
AAV82051 HCV anti  
AAV64436 Hepatitis  
ADA49671 Multi-epi  
ADO23856 Hepatitis  
ADL26465 Hepatitis  
ADL26468 Hepatitis  
AAJ01618 Hepatitis  
AAJ00864 Hepatitis  
AAJ00228 Hepatitis  
ADN64405 HLA bindi  
AAJ01478 Hepatitis

99 39 83.0 10 4 AAJ00275 Aaj00275 Hepatitis  
100 39 83.0 11 4 AAJ02942 Aaj02942 Hepatitis

## ALIGNMENTS

## RESULT 1

AAJ82996  
ID AAY82996 standard; peptide; 9 AA.

XX AC AAY82996;

XX DT 04-JUL-2000 (first entry)

XX Substituted hepatitis C virus core polypeptide antigenic fragment.

XX Hepatitis C virus; HCV; core polypeptide; antigen; epitope; agretope;  
XX anchor residue; cytotoxic T lymphocyte; CTL; immune response; MHC;  
XX Major histocompatibility complex; vaccine; treatment.

XX OS Synthetic.

XX OS Hepatitis C virus; (HCV).

XX FH Key Location/Qualifiers

XX FT Misc-difference 8 /label= Ala substituted for Leu

XX WT WO200011186-A1.

XX PN 02-MAR-2000.

XX PF 17-AUG-1999; 99WO-US018674.

XX PR 21-AUG-1998; 98US-0097446P.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Berzofsky JA, Sarobe P, Pendleton CD, Feinstein SM, Arichi T;  
XX PI Major ME;

XX DR WPI; 2000-246569/21.

XX Hepatitis C virus polypeptides is useful as a vaccine for treating

XX Hepatitis C virus infection and for activating cytotoxic T lymphocytes.

XX PS Claim 16; Page 71; 78pp; English.

XX CC Peptide fragments of the Hepatitis C virus (HCV) core polypeptide can be  
XX used to elicit an immune response. A cytotoxic T-lymphocyte (CTL)  
XX response is present in patients acutely and chronically infected by HCV.  
XX CTL's recognise viral antigens as peptides presented by Class I molecules  
XX of the major histocompatibility complex (MHC). These peptide antigens are  
XX typically 8-10 amino acids long. These peptide antigens typically have a  
XX sequence pattern or motif depending upon which particular MHC antigen  
XX they are bound and presented by. The motifs are based on the presence in  
XX precise positions in the peptide sequence of several amino acids  
XX (agretopic residues) called anchor residues, responsible for interactions  
XX between peptide and MHC molecule, as well as other secondary position  
XX that may help to stabilise the interaction. Thus, single amino acid  
XX changes within the peptide sequence can enhance the immune response.  
XX Peptides derived from the HCV core polypeptide having substituted amino  
XX acids within their sequence, specifically a hepatitis C virus core  
XX polypeptide comprising an L to an A substitution at amino acid position  
XX 139 enhance the immune response against HCV. They can therefore be used  
XX as a vaccine or for treating HCV infection. This peptide is a fragment of  
XX the substituted HCV core polypeptide and comprises the region of amino  
XX acids within that polypeptide from position 132 to position 140 which  
XX contains the amino acid substitution

XX SQ Sequence 9 AA;

Query Match . . . 100.0%; Score 47; DB 3; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
| | | | | | | |  
DB 1 DLMGYIPAV 9

## RESULT 2

AAJ82052  
ID AAB82052 standard; peptide; 9 AA.

XX AC AAB82052;

XX DT 06-AUG-2003 (revised)

XX DT 22-JUN-2001 (first entry)

XX DE HCV antigen, core 132-140.

XX Antigen; immunostimulant; vaccine; pharmaceutical composition; antiviral;  
XX viral infection.

XX OS Hepatitis C virus.

XX PN WO200124822-A2.

XX PD 12-APR-2001.

XX PF 02-OCT-2000; 2000WO-EP009657.

XX PR 01-OCT-1999; 99AT-00001680.

XX PA (CIST-) CISTEM BIOTECHNOLOGIES GMBH.

XX PI Fleitmann J, Mattner F, Buschle M, Melling J;

XX DR WPI; 2001-290577/30.

XX PT New pharmaceutical composition comprising an antigen, an  
XX immunostimulating substance and a polycationic polymer, useful in  
XX manufacturing vaccines.

XX PS Claim 12; Page 16; 20pp; English.

XX CC The present invention relates to a pharmaceutical composition comprising  
XX (a) an antigen; (b) an immunostimulating substance consisting of  
XX neuroactive compounds, hormones, compounds having growth hormone activity  
XX or their mixtures; and (c) a polycationic polymer. The present sequence  
XX is an antigenic peptide derived from Hepatitis C virus, which was used in  
XX the present invention. The composition is useful in manufacturing  
XX vaccines. (Updated on 06-AUG-2003 to correct OS field.)

XX SQ Sequence 9 AA;

Query Match . . . 100.0%; Score 47; DB 4; Length 9;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
| | | | | | | |  
DB 1 DLMGYIPAV 9

## RESULT 3

ADL25899  
ID ADL25899 standard; peptide; 15 AA.

XX AC ADL25899;

XX DT 17-JUN-2004 (first entry)

XX DE Synthetic peptide A86 derived from a conserved region of HCV.

KW HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;  
 KW major histocompatibility complex; human leukocyte antigen.  
 XX  
 XX Synthetic.

XX WO2004024182-A2.

XX 25-MAR-2004.

XX 27-AUG-2003; 2003WO-EP009482.

XX 13-SEP-2002; 2002AT-00001376.

XX 27-FEB-2003; 2003WO-EP002005.

XX 11-JUL-2003; 2003EP-00450171.

XX (INTE-) INTERCELL AG.

XX Buschle M, Habel A, Klade C, Mattner F, Otava O, Vytvytska O;

XX Zauner W, Zinke S, Kiriappos H;

XX WPI; 2004-269899/25.

XX Isolating Hepatitis C Virus peptides (HPs) which have a binding capacity  
 PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the  
 PT molecule by separating the complex from the HCV-peptides which do not  
 PT bind to the molecule.

XX Example 1; Page 30; 73pp; English.

XX The invention relates to a novel method for isolating Hepatitis C Virus  
 CC (HCV) peptides (HPs). The method of the invention has virucide activity,  
 CC and may be useful in producing a vaccine. The method is useful for  
 CC isolating Hepatitis C Virus peptides (HPs) which have a binding capacity  
 CC to a MHC/HLA molecule or a complex comprising the HCV-peptide and the  
 CC MHC/HLA molecule for preparing a vaccine against HCV infection. The T  
 CC cells, a T cell clone or a T cell population or preparation is useful for  
 CC identifying heteroclitic epitopes or for preparing a composition for  
 CC treating HCV infection. The present sequence represents a synthetic  
 CC peptide derived from a conserved region of HCV.

XX Sequence 15 AA;

Query Match 91.5%; Score 43; DB 8; Length 15;  
 Best Local Similarity 88.9%; Pred. No. 0.095;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9

DB 7 DLMGYIPW 15

RESULT 4

ADL25904

ID ADL25904 standard; peptide; 15 AA.

XX ADL25904;

XX 17-JUN-2004 (first entry)

XX Synthetic peptide A91 derived from a conserved region of HCV.

XX HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;

XX major histocompatibility complex; human leukocyte antigen.

XX Synthetic.

XX WO2004024182-A2.

XX 25-MAR-2004.

XX 27-AUG-2003; 2003WO-EP009482.

XX 13-SEP-2002; 2002AT-00001376.

PR 27-FEB-2003; 2003WO-EP002005.

PR 11-JUL-2003; 2003EP-00450171.

XX (INTE-) INTERCELL AG.

XX Buschle M, Habel A, Klade C, Mattner F, Otava O, Vytvytska O;

XX Zauner W, Zinke S, Kiriappos H;

XX WPI; 2004-269899/25.

XX Isolating Hepatitis C Virus peptides (HPs) which have a binding capacity  
 PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the  
 PT molecule by separating the complex from the HCV-peptides which do not  
 PT bind to the molecule.

XX Example 1; Page 30; 73pp; English.

XX The invention relates to a novel method for isolating Hepatitis C Virus  
 CC (HCV) peptides (HPs). The method of the invention has virucide activity,  
 CC and may be useful in producing a vaccine. The method is useful for  
 CC isolating Hepatitis C Virus peptides (HPs) which have a binding capacity  
 CC to a MHC/HLA molecule or a complex comprising the HCV-peptide and the  
 CC MHC/HLA molecule for preparing a vaccine against HCV infection. The T  
 CC cells, a T cell clone or a T cell population or preparation is useful for  
 CC identifying heteroclitic epitopes or for preparing a composition for  
 CC treating HCV infection. The present sequence represents a synthetic  
 CC peptide derived from a conserved region of HCV.

XX Sequence 15 AA;

Query Match 91.5%; Score 43; DB 8; Length 15;  
 Best Local Similarity 88.9%; Pred. No. 0.095;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9

DB 2 DLMGYIPV 10

RESULT 5

ADL25905

ID ADL25905 standard; peptide; 15 AA.

XX ADL25905;

XX 17-JUN-2004 (first entry)

XX Synthetic peptide A92 derived from a conserved region of HCV.

XX HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;

XX major histocompatibility complex; human leukocyte antigen.

XX Synthetic.

XX WO2004024182-A2.

XX 25-MAR-2004.

XX 27-AUG-2003; 2003WO-EP009482.

XX 13-SEP-2002; 2002AT-00001376.

XX 27-FEB-2003; 2003WO-EP002005.

XX 11-JUL-2003; 2003EP-00450171.

XX (INTE-) INTERCELL AG.

XX Buschle M, Habel A, Klade C, Mattner F, Otava O, Vytvytska O;

XX Zauner W, Zinke S, Kiriappos H;

XX WPI; 2004-269899/25.

XX Isolating Hepatitis C Virus peptides (HPs) which have a binding capacity  
 PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the

PT molecule by separating the complex from the HCV-peptides which do not bind to the molecule.

XX

PS Example 1: Page 30; 73pp; English.

The invention relates to a novel method for isolating Hepatitis C Virus (HCV) peptides (HPs). The method of the invention has virucide activity, and may be useful in producing a vaccine. The method is useful for isolating Hepatitis C Virus peptides (HPs) which have a binding capacity to a MHC/HLA molecule or a complex comprising the HCV-peptide and the MHC/HLA molecule for preparing a vaccine against HCV infection. The cells, a T cell clone or a T cell population or preparation is useful for identifying heteroclitic epitopes or for preparing a composition for treating HCV infection. The present sequence represents a synthetic peptide derived from a conserved region of HCV.

Sequence 15 AA:

Query Match 91.5%; Score 43; DB 8; Length 15;  
Best Local Similarity 88.9%; Pred. No. 0.095;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | |  
pb 1 DLMGYIPVV 9

RESIST 6

ADL25900  
ID ADL25900 standard; peptide; 15 AA.

AC ADL25900:

XX  
XX

DT 17-JUN-2004 (first entry)

DE Synthetic peptide A87 derived from a conserved region of HCV.

**XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX**

KW HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;  
KW major histocompatibility complex; human leukocyte antigen.

OS Synthetic.

XX PN WO2004024182-A2.

XX  
PD 25-MAR-2004.

XX  
PF 27-AUG-2003; 2003WO-EP009482.

XX

PR 13-SEP-2002; 2002AT-00001376.

PR 27-FEB-2003; 2003WO-EP002005.  
PR 11-JUL-2003; 2003EP-00450171.

PA (INTE-) INTERCELL AG.

XX  
PI Buschle M, Habel A, Klade C, Mattner F, Otava O, Vytvytska O;  
PI Zauner W, Zinke S, Kirlappos H;

XX  
DR WPI; 2004-269899/25.

XX Isolating Hepatitis C Virus peptides (HPs) which have a binding capacity  
PT to a MHC/HLA molecule or a complex comprising the HCV-peptide and the  
PT molecule by separating the complex from the HCV-peptides which do not  
PT bind to the molecule.  
PT

Example 1: Page 30: 73pp; English.

The invention relates to a novel method for isolating Hepatitis C Virus (HCV) peptides (Hps). The method of the invention has virucide activity, and may be useful in producing a vaccine. The method is useful for isolating Hepatitis C Virus peptides (Hps) which have a binding capacity to a MHC/HLA molecule or a complex comprising the HCV-peptide and the MHC/HLA molecule for preparing a vaccine against HCV infection. The T cells, a T cell clone or a T cell population or preparation is useful for

XX	HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;
KW	major histocompatibility complex; human leukocyte antigen.
XX	Synthetic.
PN	WO2004024182-A2.
XX	
PD	25-MAR-2004.
XX	
XX	27-AUG-2003; 2003WO-EP009482.
XX	
PR	13-SEP-2002; 2002AT-00001376.
PR	27-FEB-2003; 2003WO-EP002005.
PR	11-JUL-2003; 2003BP-00450171.
XX	
PA	(INTE-) INTERCELL AG.
PI	Buschle M, Habel A, Klade C, Mattner F, Otava O, Vytvytska O;
PI	Zauner W, Zinke S, Kiriappos H;
XX	
DR	WPI; 2004-269899/25.
XX	
PT	Isolating Hepatitis C Virus peptides (Hps) which have a binding capacity
PT	to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
PT	molecule by separating the complex from the HCV-peptides which do not
PT	bind to the molecule.
XX	
PS	Example 1; Page 30; 73pp; English.
XX	
CC	The invention relates to a novel method for isolating Hepatitis C Virus
CC	(HCV) peptides (Hps). The method of the invention has virucide activity,
CC	and may be useful in producing a vaccine. The method is useful for
CC	isolating Hepatitis C Virus peptides (Hps) which have a binding capacity
CC	to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
CC	MHC/HLA molecule for preparing a vaccine against HCV infection. The T
CC	cells, a T cell clone or a T cell population or preparation is useful for
CC	identifying heteroclitic epitopes or for preparing a composition for
CC	treating HCV infection. The present sequence represents a synthetic
CC	peptide derived from a conserved region of HCV.
XX	
SQ	Sequence 15 AA;
	Query Match 91.5%; Score 43; DB 8; Length 15;
	Best Local Similarity 88.9%; Pred. No. 0.095;
	Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0
QY	1 DLMGYIPAV 9
Db	4 DLMGYIPW 12
	RESULT 10
	ADL26466
ID	ADL26466 standard; peptide; 36 AA.
XX	
AC	ADL26466;
XX	
DT	17-JUN-2004 (first entry)
XX	
DE	Synthetic peptide 1649 derived from a conserved region of HCV.
XX	
KW	HCV; hepatitis C virus; virucide; vaccine; MHC; HLA;
KW	major histocompatibility complex; human leukocyte antigen.
XX	Synthetic.
OS	
XX	
PN	WO2004024182-A2.
XX	
PD	25-MAR-2004.
XX	
XX	27-AUG-2003; 2003WO-EP009482.
XX	

```

PR 13-SEP-2002; 2002AT-00001376.
PR 27-FEB-2003; 2003WO-EP002005.
PR 11-JUL-2003; 2003EP-00450171.
XX (INTE-) INTERCELL AG.
XX
XX Buechle M, Habel A, Klade C, Mattner F, Otava O, Vytvytska O;
XX Zauner W, Zinke S, Kirlappos H;
XX WPI; 2004-269899/25.
XX
XX Isolating Hepatitis C Virus peptides (HVs) which have a binding capacity
XX to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
XX molecule by separating the complex from the HCV-peptides which do not
XX bind to the molecule.
XX
XX Example 1; Page 32; 73pp; English.
XX
XX The invention relates to a novel method for isolating Hepatitis C Virus
XX (HCV) peptides (HVs). The method of the invention has virucide activity,
XX and may be useful in producing a vaccine. The method is useful for
XX isolating Hepatitis C Virus peptides (HVs) which have a binding capacity
XX to a MHC/HLA molecule or a complex comprising the HCV-peptide and the
XX MHC/HLA molecule for preparing a vaccine against HCV infection. The T
XX cells, a T cell clone or a T cell population or preparation is useful for
XX identifying heteroclitic epitopes or for preparing a composition for
XX treating HCV infection. The present sequence represents a synthetic
XX peptide derived from a conserved region of HCV.
XX
XX Sequence 36 AA;
XX
XX Query Match 91.5%; Score 43; DB 8; Length 36;
XX Best Local Similarity 88.9%; Pred. No. 0.25;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 DLMGYIPAV 9
XX |||||
XX 22 DLMGYIPV 30
XX
XX RESULT 11
XX AAR73111
XX ID AAR73111 standard; peptide; 9 AA.
XX AC AAR73111;
XX
XX 25-MAR-2003 (revised)
XX 16-JUN-1995 (first entry)
XX
XX Antigen fragment 8 from HCV has binding affinity for HLA-2.1.
XX
XX antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; EBV; HIV1;
XX plasma specific antigen; hepatitis B virus; Epstein Barr;
XX human immunodeficiency virus; human papilloma virus; p53; c-ERB2; MAGE-1;
XX melanoma antigen-1; core antigen; surface antigen;
XX pharmaceutical composition; in vivo; ex vivo; therapeutic; diagnostic;
XX MHC class I molecule; major histocompatibility complex; HLA-A2.1; 9mer;
XX 10mer; anchor; human leukocyte antigen; PLP; 8mer; algorithm prediction;
XX MBP; CMV; cytomegalovirus; HSV; herpes simplex virus; influenza A; M1.
XX
XX Hepatitis C virus.
XX
XX WO9420127-A1.
XX
XX 15-SEP-1994.
XX
XX 04-MAR-1994; 94WO-US002353.
XX
XX 05-MAR-1993; 93US-00027146.
XX 04-JUN-1993; 93US-00073205.
XX 29-NOV-1993; 93US-00159184.
XX
XX (CYTE-) CYTEL CORP.
XX
XX Grey HM, Sette A, Sidney J, Kast W;
XX WPI; 1994-302678/37.
XX
XX Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for
XX treatment or prophylaxis of cancer, virus infection or autoimmune
XX diseases.
XX
XX Disclosure; Page 90; 138pp; English.
XX
XX AAR73058-121 are potential peptide binders of HLA-A2.1 motif. Using
XX motifs disclosed in the invention, these peptides were screened for
XX further motifs. Only peptides with binding affinity of at least 1%
XX (binding affinity is expressed as an IC50 value) as compared to the
XX standard peptide (AAR71293) in assays. This peptide from hepatitis C
XX virus has a binding value of 0.0550. The peptides of the invention can
XX induce cytotoxic T lymphocytes which can react with target cells. They
XX can be used for the treatment or prophylaxis of cancer, eg, prostate
XX cancer or lymphoma, etc. (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 9 AA;
XX
XX Query Match 89.4%; Score 42; DB 2; Length 9;
XX Best Local Similarity 88.9%; Pred. No. 1.8e+06;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 DLMGYIPAV 9
XX |||||
XX 1 DLMGYIPLV 9
XX
XX RESULT 12
XX AAR59125
XX ID AAR59125 standard; peptide; 9 AA.
XX AC AAR59125;
XX
XX 16-OCT-2003 (revised)
XX 25-MAR-2003 (revised)
XX 02-MAY-1995 (first entry)
XX
XX Peptide fragment (1.0816) of HCV binds HLA-A2.1.
XX
XX antigen; epitope; immunogenic target protein; PSA; HBVc; HBVs; EBV; HIV1;
XX core antigen; surface antigen; pharmaceutical composition; in vivo;
XX ex vivo; therapeutic; diagnostic; MHC class I molecule;
XX major histocompatibility complex; HLA-A2.1; 9mer; 10mer; anchor;
XX human leukocyte antigen.
XX
XX Hepatitis C virus; strain.
XX
XX WO9420127-A1.
XX
XX 15-SEP-1994.
XX
XX 04-MAR-1994; 94WO-US002353.
XX
XX 05-MAR-1993; 93US-00027146.
XX 04-JUN-1993; 93US-00073205.
XX 29-NOV-1993; 93US-00159184.
XX
XX (CYTE-) CYTEL CORP.
XX
XX Grey HM, Sette A, Sidney J, Kast W;
XX WPI; 1994-302678/37.
XX
XX Immunogenic peptide(s) having an HLA-A2.1 binding motif - used for
XX treatment or prophylaxis of cancer, virus infection or autoimmune
XX diseases.
XX
XX Example 5; Page 100; 138pp; English.

```



CC AAR59104-264 are immunogenic 9mer peptides that contain a HLA-A2.1  
 CC binding motif. These peptides bind HLA-A2.1 and have a binding affinity  
 CC of at least 1% as compared to a reference peptide (AAR71293). AAR59125  
 CC has an IC50 of 0.055 and the sequence occurs at position 132 in the HCV  
 CC CORE protein. The peptides of the invention can induce cytotoxic T  
 CC lymphocytes which can react with target cells. They can be used for the  
 CC treatment or prophylaxis of cancer, eg. prostate cancer or lymphoma, etc.  
 CC (Updated on 25-MAR-2003 to correct PN field.) (Updated on 16-OCT-2003 to  
 CC standardise OS field)  
 CC  
 SQ Sequence 9 AA;  
 Query Match 89.4%; Score 42; DB 2; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DLMGYIPAV 9  
 Db |||||  
 1 DLMGYIPLV 9  
 RESULT 13  
 AAR78946  
 ID AAR78946 standard; peptide; 9 AA.  
 XX  
 AC AAR78946;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 01-APR-1996 (first entry)  
 XX  
 DE HCV core 132-140 cytotoxic T lymphocyte epitope.  
 XX  
 KW HCV core 132-140; cytotoxic T; CTL; epitope; helper T; HTL; cell;  
 KW lymphocyte; antigens; treatment; disease prevention; hepatitis C; non-A;  
 KW non-B.  
 XX  
 OS Hepatitis C virus.  
 XX  
 PN WO9522317-A1.  
 XX  
 PD 24-AUG-1995.  
 XX  
 PF 16-FEB-1995; 95WO-US002121.  
 XX  
 PR 16-FEB-1994; 94US-00197484.  
 XX  
 PA (CYTE-) CYTEL CORP.  
 XX  
 PI Vitiello MA, Chesnut RW, Sette AD, Celis E, Grey H;  
 XX  
 DR WPI; 1995-302545/39.  
 XX  
 PT Compn. inducing cytotoxic T lymphocyte response to pref. viral,  
 PT bacterial, parasitic or tumour antigens - useful in the treatment and  
 PT prevention of diseases associated with the antigen e.g. hepatitis B.  
 XX  
 PS Example 12; Page 70; 109pp; English.  
 XX  
 CC A compen. which induces a cytotoxic T lymphocyte (CTL) response to a  
 CC hepatitis C virus (HCV) antigen (Ag) in a mammal comprises, a HCV CTL Ag  
 CC response inducing peptide (i.e. AAR78941-R78955) and a lipid conjugated  
 CC helper T cell inducing peptide. The compen. is useful in the treatment  
 CC and prevention of hepatitis C. (Updated on 25-MAR-2003 to correct PI  
 CC field.)  
 XX  
 SQ Sequence 9 AA;  
 Query Match 89.4%; Score 42; DB 2; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DLMGYIPAV 9

Db |||||  
 1 DLMGYIPLV 9  
 RESULT 14  
 AAR87652  
 ID AAR87652 standard; peptide; 9 AA.  
 XX  
 AC AAR87652;  
 XX  
 DT 20-MAY-1996 (first entry)  
 XX  
 DE Hepatitis C virus core peptide used for treating HCV infection.  
 XX  
 KW Hepatitis C virus; cytotoxic T lymphocyte; vaccine; immunisation;  
 KW diagnosis; HCV; CTL; core; HLA.  
 XX  
 OS Hepatitis C virus.  
 XX  
 PN WO9527733-A1.  
 XX  
 PD 19-OCT-1995.  
 XX  
 PF 07-APR-1995; 95WO-US003935.  
 XX  
 PR 08-APR-1994; 94US-00224973.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Berzofsky JA, Feinstone S, Shirai M;  
 XX  
 DR WPI; 1995-366355/47.  
 XX  
 PT Hepatitis C virus core peptide(s) stimulate cytotoxic T lymphocyte  
 PT response - used for prevention, treatment or diagnosis of HCV infection.  
 XX  
 PS Claim 6; Page 45; 58pp; English.  
 XX  
 CC AAR87651-53 are hepatitis C virus (HCV) core region peptides which can be  
 CC used for the immunisation of an individual against HCV. They can also be  
 CC used for diagnosing exposure of a patient to HCV or for predicting a  
 CC patient's clinical course following HCV infection. The peptides provoke  
 CC the activation of cytotoxic T cells which attack any cell infected with  
 CC HCV  
 XX  
 SQ Sequence 9 AA;  
 Query Match 89.4%; Score 42; DB 2; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DLMGYIPAV 9  
 Db |||||  
 1 DLMGYIPLV 9  
 RESULT 15  
 AAR91054  
 ID AAR91054 standard; peptide; 9 AA.  
 XX  
 AC AAR91054;  
 XX  
 DT 25-APR-1996 (first entry)  
 XX  
 DE HCV-1 derived peptide tested for usefulness in a HCV vaccine.  
 XX  
 KW Hepatitis C virus; HCV; epitope; vaccine; immunogen.  
 XX  
 OS Hepatitis C virus.  
 XX  
 PN WO9525122-A1.  
 XX  
 PD 21-SEP-1995.

XX PF 16-MAR-1995; 95WO-US003224.  
 XX PR 17-MAR-1994; 94US-00214650.  
 XX PA (SCRI ) SCRIPPS RES INST.  
 XX PI Chisari FV, Cerny A;  
 XX DR WPI; 1995-336941/43.  
 XX PT Novel molecule comprising a cytotoxic T cell epitope - used to vaccinate  
 XX PT against hepatitis C viral infection.  
 XX PS Example 1; Page 46; 85pp; English.  
 XX CC AAR84570-616, AAR84885-90 and AAR91054 are all HCV-1 derived peptides  
 CC from the core, E1, E2/NS1, NS2, NS3, NS4 or NS5 regions. The peptides  
 CC were tested for peptide specific cytotoxic T-cell activity. The peptides  
 CC AAR84570-77 were found to have substantial homology with a T- cell  
 CC epitope and are useful in vaccines against HCV infection  
 XX SQ Sequence 9 AA;  
 CC  
 CC Query Match 89.4%; Score 42; DB 2; Length 9;  
 CC Best Local Similarity 88.9%; Pred. NO. 1.8e+06;  
 CC Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DLMGYIPAV 9  
 DB 1 DLMGYIPLV 9  
 CC  
 CC RESULT 16  
 CC AAW39446  
 CC ID AAW39446 standard; peptide; 9 AA.  
 CC XX AAW39446;  
 CC AC AAW39446;  
 CC DT 11-JUN-1998 (first entry)  
 CC XX Human HCV1core-131 variant immunogenic peptide.  
 CC DE T cell epitope; immune response; human leukocyte antigen; HLA Class I;  
 CC XX vaccine; immunogenic; major histocompatibility complex; MHC; B cell;  
 CC KW disease; anti-tumour; anti-viral.  
 CC XX Synthetic.  
 CC OS Homo sapiens.  
 CC XX WO9741440-A1.  
 CC PN 06-NOV-1997.  
 CC PD 28-APR-1997; 97WO-NL000229.  
 CC XX 26-APR-1996; 96EP-00201145.  
 CC PR 23-DEC-1996; 96EP-00203670.  
 CC XX (UYLE-) RIJKSUNIV LEIDEN.  
 CC PA (SCIS-) SCI SEED CAPITAL INVESTMENTS BV.  
 CC PI Van Der Burg SH, Kast WM, Toes REM, Offringa R, Melief CJM;  
 CC XX WPI; 1997-549891/50.  
 CC DR Method of selecting T cell peptide epitope(s) - by measuring the  
 CC XX stability of HLA class I-peptide complexes on intact B cells.  
 CC PS Example 2; Page 27; 109pp; English.  
 CC XX Peptides AAW39430-W39734 are used in a novel method for the selection of  
 CC immunogenic T-cell peptide epitopes present in polypeptide antigens. The

CC method involves the identification of peptide sequences capable of  
 CC binding to an HLA (human leukocyte antigen) class I molecule and  
 CC measuring the binding of this epitope peptide to the HLA class I peptide.  
 CC The stability of binding of the peptide and MHC (major histocompatibility  
 CC complex) class I molecule is measured on intact human B cells carrying  
 CC the MHC molecule at their cell surfaces. The method can be used to select  
 CC peptide epitopes for generating vaccines against a disease associated  
 CC with the polypeptide, e.g. cancers or AIDS. The peptide epitopes are  
 CC especially T-cell peptide epitopes with strong anti-tumour and anti-viral  
 CC immune responses. Peptide AAW39446 is a shorter variant the HCV1core-131  
 CC peptide which is a HLA-A\*0201 restricted cytotoxic T lymphocyte (CTL)  
 CC epitope  
 XX SQ Sequence 9 AA;  
 CC  
 CC Query Match 89.4%; Score 42; DB 2; Length 9;  
 CC Best Local Similarity 88.9%; Pred. NO. 1.8e+06;  
 CC Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DLMGYIPAV 9  
 DB 1 DLMGYIPLV 9  
 CC  
 CC RESULT 17  
 CC AAW54635  
 CC ID AAW54635 standard; peptide; 9 AA.  
 CC XX AAW54635;  
 CC AC AAW54635;  
 CC DT 25-SEP-1998 (first entry)  
 CC XX Peptide from HCV core 132-140.  
 CC DE Mannose; antigen; antigen-presenting cell; mannosylated peptide; T cell;  
 CC KW vaccine; treatment.  
 CC XX Synthetic.  
 CC OS WO9813378-A1.  
 CC PN 02-APR-1998.  
 CC PD 25-SEP-1997; 97WO-NL000536.  
 CC XX 26-SEP-1996; 96EP-00202701.  
 CC PR (UYLE-) RIJKSUNIV LEIDEN.  
 CC PA Koning F, Drijfhout JW;  
 CC PI WPI; 1998-230631/20.  
 CC DR Increasing uptake and presentation of antigen(s) - by adding mannose  
 CC XX residue(s) to antigen for increasing T cell response, useful in, e.g.  
 CC PT vaccines against viral infection(s).  
 CC PS Disclosure; Page 28; 47pp; English.  
 CC XX The peptides AAW54559-W54809 are examples of peptides to which at least 1  
 CC (preferably 2) mannose can be attached to increase their uptake as  
 CC antigens by antigen-presenting cells. Uptake of agonist mannosylated  
 CC peptides will increase the T cell response, whereas uptake of antagonist  
 CC peptides blocks the T cell response. Blocking binding of immunogenic  
 CC autoantigens can be used in treatment of type I diabetes, rheumatoid  
 CC arthritis, graft rejection etc., also to induce T-cell non-  
 CC responsiveness. Vaccines containing mannosylated antigen are used to  
 CC prevent or treat infections by, e.g. bacteria, viruses, fungi, helminths  
 CC and parasites  
 XX SQ Sequence 9 AA;  
 CC  
 CC Query Match 89.4%; Score 42; DB 2; Length 9;  
 CC

Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | |  
Db 1 DLMGYIPLV 9

## RESULT 18

AAV10219

ID AAY10219 standard; peptide; 9 AA.

XX AC AAY10219;

DT 12-MAY-1999 (first entry)

XX T cell epitope/MHC ligand SEQ ID NO:149.

XX Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;

XX immunisation; tumour; infectious disease; immunotherapy; cancer;

XX malignant melanoma; viral disease; hepatitis; AIDS.

XX Synthetic.

XX Hepatitis C virus.

XX WO9902183-A2.

XX 21-JAN-1999.

XX 10-JUL-1998; 98WO-US014289.

XX 10-JUL-1997; 97CA-02209815.

XX 10-DEC-1997; 97US-00988320.

XX (CTL1-) CTL IMMUNOTHERAPIES CORP.

XX Kuendig TM, Simard JJJ;

XX WPI; 1999-120514/10.

XX Inducing a cytotoxic T lymphocyte response - by maintaining a level of

XX antigen in the lymphatic system of a mammal so as to provide a sustained

XX CTL response, used to treat, e.g. AIDS.

XX Disclosure; Page 29; 199pp; English.

XX The present invention describes a method of inducing and/or sustaining an

XX immunological cytotoxic T lymphocyte (CTL) response in a mammal. The

XX method comprises: (a) delivering an antigen to the mammal at a level to

XX induce an immunological CTL response in the mammal; and (b) maintaining

XX the level of the antigen in the mammal's lymphatic system to maintain the

XX immunologic CTL response. The method can be used for the delivery of e.g.

XX a differentiation antigen, a tumour-specific multilineage antigen, an

XX embryonic antigen, an oncogene antigen, a mutated tumour-suppressor gene

XX antigen, or a viral antigen. They can be used for the treatment of

XX disease such as cancer, e.g. malignant melanoma or infectious disease,

XX e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery

XX to the lymphatic system provides for potent CTL stimulation that takes

XX place in the milieu of the lymphoid organ, and it sustains stimulation

XX that is necessary to keep CTL active, cytotoxic and recirculating through

XX the body. AAY10071 to AAY10639 represent examples of peptide antigens

XX given in the present invention

XX Sequence 9 AA;

XX Query Match 89.4%; Score 42; DB 2; Length 9;

XX Best Local Similarity 88.9%; Pred. No. 1.8e+06;

XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9

| | | | | | |

Db 1 DLMGYIPLV 9

| | | | | | |

RESULT 20

AAV73106

ID AAY73106 standard; peptide; 9 AA.

XX AC AAY73106;

XX AAY73106;

## RESULT 19

AAV10441

ID AAY10441 standard; peptide; 9 AA.

XX AC AAY10441;

DT 12-MAY-1999 (first entry)

XX HLA Class I motif peptide SEQ ID NO:371.

XX Cytotoxic T-lymphocyte response; CTL; antigen; lymphatic system;

XX immunisation; tumour; infectious disease; immunotherapy; cancer;

XX malignant melanoma; viral disease; hepatitis; AIDS.

XX Synthetic.

XX Human papillomavirus.

XX WO9902183-A2.

XX 21-JAN-1999.

XX 10-JUL-1998; 98WO-US014289.

XX 10-JUL-1997; 97CA-02209815.

XX 10-DEC-1997; 97US-00988320.

XX (CTL1-) CTL IMMUNOTHERAPIES CORP.

XX Kuendig TM, Simard JJJ;

XX WPI; 1999-120514/10.

XX Inducing a cytotoxic T lymphocyte response - by maintaining a level of

XX antigen in the lymphatic system of a mammal so as to provide a sustained

XX CTL response, used to treat, e.g. AIDS.

XX Disclosure; Page 40; 199pp; English.

XX The present invention describes a method of inducing and/or sustaining an

XX immunological cytotoxic T lymphocyte (CTL) response in a mammal. The

XX method comprises: (a) delivering an antigen to the mammal at a level to

XX induce an immunological CTL response in the mammal; and (b) maintaining

XX the level of the antigen in the mammal's lymphatic system to maintain the

XX immunologic CTL response. The method can be used for the delivery of e.g.

XX a differentiation antigen, a tumour-specific multilineage antigen, an

XX embryonic antigen, an oncogene antigen, a mutated tumour-suppressor gene

XX antigen, or a viral antigen. They can be used for the treatment of

XX disease such as cancer, e.g. malignant melanoma or infectious disease,

XX e.g. viral disease such as hepatitis or AIDS. Sustained antigen delivery

XX to the lymphatic system provides for potent CTL stimulation that takes

XX place in the milieu of the lymphoid organ, and it sustains stimulation

XX that is necessary to keep CTL active, cytotoxic and recirculating through

XX the body. AAY10071 to AAY10639 represent examples of peptide antigens

XX given in the present invention

XX Sequence 9 AA;

XX Query Match 89.4%; Score 42; DB 2; Length 9;

XX Best Local Similarity 88.9%; Pred. No. 1.8e+06;

XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9

| | | | | | |

Db 1 DLMGYIPLV 9

| | | | | | |

```

DT 06-AUG-2003 (revised)
DT 28-FEB-2000 (first entry)
DE Hepatitis C virus (HCV)-derived MHC class I (CTL) epitope, #264.
DE
XX
XX Chimeric; pan DR epitope; expression vector; promoter;
XX major histocompatibility complex; MHC; targeting; peptide; epitope;
XX antigen; presentation; class I; cytosolic pathway; endoplasmic reticulum;
XX class II; extracellular antigen; endocytic pathway; helper T lymphocyte;
XX HTL; universal epitope; cytotoxic T lymphocyte; CTL; immune response;
XX immunogenicity; assay; vaccine; immunity; infection; pathogen; virus;
XX HIV; HBV; HCV; hepatitis B; hepatitis C; bacterium; protozoan;
XX tumour cell; autoimmune disease; activation; antiviral; antimalarial;
XX immunoprotective.
XX
XX Synthetic.
XX
XX Hepatitis C virus.
XX
XX WO9958658-A2.
XX
XX 18-NOV-1999.
XX
XX 13-MAY-1999; 99WO-US010646.
XX
XX 13-MAY-1998; 98US-00078904.
XX
XX 15-MAY-1998; 98US-0085751P.
XX
XX (EPIM-) EPIMUNE INC.
XX
XX Fikes JD, Hermanson GG, Sette A, Ishioka GY, Livingston B;
XX Chesnut RW;
XX
XX WPI; 2000-039103/03.
XX
XX Expression vectors encoding major histocompatibility targeting sequence,
XX used as, e.g. tumor vaccines.
XX
XX Claim 11; Page 68; 130pp; English.
XX
XX Sequences AAY73103-Y73145 represent hepatitis C virus (HCV)-derived MHC
XX class I (CTL) epitopes which are claimed for use in the present
XX invention. The invention relates to a novel expression vector comprising
XX a promoter operably linked to a fusion gene encoding a major
XX histocompatibility complex (MHC) targeting sequence, and two or more
XX heterologous peptide epitopes. The MHC targeting sequence may be a class
XX I targeting sequence, which directs an MHC class I epitope to a
XX cytosolic pathway or to the endoplasmic reticulum, or an MHC class II
XX targeting sequence, which directs extracellular antigens to enter the
XX endocytic pathway to be processed into antigen peptides for presentation
XX on MHC class II molecules. The heterologous epitopes may comprise either
XX helper T lymphocyte (HTL) epitopes, or a cytotoxic T lymphocyte (CTL)
XX epitope and a universal HTL epitope such as a pan DR epitope (PADRE). The
XX vectors are useful for stimulating an immune response in vivo, as well as
XX for use in assaying the human immunogenicity of a human T cell peptide
XX epitope in vivo in a non-human mammal. They provide a nucleic acid
XX vaccine for enhancing immunity against infectious pathogens, such as
XX viruses (e.g., HIV, hepatitis B (HBV) and hepatitis C (HCV)), bacteria,
XX protozoa (e.g., Plasmodium falciparum, the cause of malaria) and also
XX tumour cells and autoimmune diseases. Universal MHC class II epitopes are
XX advantageously combined with other MHC class I and class II epitopes to
XX increase the number of cells that are activated in response to a given
XX antigen and provide a broader population coverage of MHC-reactive
XX alleles. (Updated on 06-AUG-2003 to correct OS field.)
XX
XX Sequence 9 AA;
XX
XX Query Match 89.4%; Score 42; DB 3; Length 9;
XX Best Local Similarity 88.9%; Pred. No. 1.8e+06;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 DLMGYIPAV 9
XX Db 1 DLMGYIPLV 9
XX
XX RESULT 22

```

---

```

RESULT 21
AAY82998
ID AAY82998 standard; peptide; 9 AA.
XX
XX AAY82998;
AC
XX 12-SEP-2003 (revised)
XX 04-JUL-2000 (first entry)
DT
XX Hepatitis C virus core polypeptide antigenic fragment.
DE
XX Hepatitis C virus; HCV; core polypeptide; antigen; epitope; agretope;
XX anchor residue; cytotoxic T lymphocyte; CTL; immune response; MHC;
XX Major histocompatibility complex; vaccine; treatment.
XX
XX Hepatitis C virus; (HCV).
OS
XX WO200011186-A1.
XX
XX 02-MAR-2000.
XX
XX 17-AUG-1999; 99WO-US018674.
XX
XX 21-AUG-1998; 98US-0097446P.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Berzofsky JA, Sarobe P, Pendleton CD, Feinstein SM, Arichi T;
XX Major ME;
XX
XX WPI; 2000-246569/21.
XX
XX Hepatitis C virus polypeptides is useful as a vaccine for treating
XX Hepatitis C virus infection and for activating cytotoxic T lymphocytes.
XX
XX Claim 33; Page 72; 78pp; English.
XX
XX Peptide fragments of the Hepatitis C virus (HCV) core polypeptide can be
XX used to elicit an immune response. A cytotoxic T-lymphocyte (CTL)
XX response is present in patients acutely and chronically infected by HCV.
XX CTL's recognise viral antigens as peptides presented by Class I molecules
XX of the major histocompatibility complex (MHC). These peptide antigens are
XX typically 8-10 amino acids long. These peptide antigens typically have a
XX sequence pattern or motif depending upon which particular MHC antigen
XX they are bound and presented by. The motifs are based on the presence in
XX precise positions in the peptide sequence of several amino acids
XX (agretopic residues) called anchor residues, responsible for interactions
XX between peptide and MHC molecule, as well as other secondary position
XX that may help to stabilise the interaction. Thus, single amino acid
XX changes within the peptide sequence can enhance the immune response.
XX Peptides derived from the HCV core polypeptide having substituted amino
XX acids within their sequence, specifically a hepatitis C virus core
XX polypeptide comprising an L to an A substitution at amino acid position
XX 139 enhance the immune response against HCV. They can therefore be used
XX as a vaccine or for treating HCV infection. This peptide is derived from
XX the wild type HCV core polypeptide and comprises the region of amino
XX acids within that polypeptide from position 132 to position 140. (Updated
XX on 12-SEP-2003 to standardise OS field)
XX
XX Sequence 9 AA;
XX
XX Query Match 89.4%; Score 42; DB 3; Length 9;
XX Best Local Similarity 88.9%; Pred. No. 1.8e+06;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 DLMGYIPAV 9
XX Db 1 DLMGYIPLV 9
XX
XX RESULT 22

```

```
AAJ03866
ID AAJ03866 standard; peptide; 9 AA.
AC AAJ03866;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #3857.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus.
XX
PS Disclosure; Page 144; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
SQ Sequence 9 AA;
    Query Match      89.4%; Score 42; DB 4; Length 9;
    Best Local Similarity 88.9%; Pred. No. 1.8e+06;
    Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9
DB |||||
  1 DLMGYIPLV 9

RESULT 24
AAJ03964
ID AAJ03964 standard; peptide; 9 AA.
XX
AC AAJ03964;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #3955.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX
DR WPI; 2001-308046/32.
XX
PT A new composition useful as a vaccines against hepatitis C virus.
XX
PS Example 3; Page 196; 214pp; English.
XX
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention
XX
SQ Sequence 9 AA;
    Query Match      89.4%; Score 42; DB 4; Length 9;
    Best Local Similarity 88.9%; Pred. No. 1.8e+06;
    Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

AAJ01746
ID AAJ01746 standard; peptide; 9 AA.
XX
AC AAJ01746;
XX
DT 02-JUL-2001 (first entry)
XX
DE Hepatitis C virus epitope #1737.
XX
KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX
OS Hepatitis C virus.
XX
PN WO200121189-A1.
XX
PD 29-MAR-2001.
XX
PF 19-JUL-2000; 2000WO-US019774.
XX
PR 19-JUL-1999; 99US-00357737.
XX
```

```

Qy      1 DLMGYIPAV 9
Db      1 DLMGYIPLV 9

RESULT 25
AAJ00011
ID AAJ00011 standard; peptide; 9 AA.
XX AC
XX AAJ00011;
DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #2.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX SQ Sequence 9 AA;

Query Match      89.4%; Score 42; DB 4; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.8e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      1 DLMGYIPAV 9
Db      1 DLMGYIPLV 9

RESULT 27
AAJ03787
ID AAJ03787 standard; peptide; 9 AA.
XX AC
XX AAJ03787;
DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #3778.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US019774.
XX PR 19-JUL-1999; 99US-00357737.
XX PA (EPIM-) EPIMMUNE INC.
XX PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
XX PI Baker DM, Cellis E, Kubo RT, Grey HM;
XX DR WPI; 2001-308046/32.
XX SQ A new composition useful as a vaccines against hepatitis C virus.
Disclosure; Page 106; 214pp; English.
The present invention describes a composition comprising a prepared hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121. These are derived from HCV HLA-binding motifs. They are useful in vaccines for the prevention and treatment of HCV infection in humans. The present sequence is an epitope used in the disclosure of the invention

Qy      1 DLMGYIPAV 9
Db      1 DLMGYIPLV 9

RESULT 26
AAJ00229
ID AAJ00229 standard; peptide; 9 AA.
XX AC
XX AAJ00229;
DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #220.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX OS Hepatitis C virus.
XX PN WO200121189-A1.

```

CC present sequence is an epitope used in the disclosure of the invention  
XX  
SQ Sequence 9 AA;

Query Match 89.4%; Score 42; DB 4; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | |  
Db 1 DLMGYIPLV 9

RESULT 28  
AAJ04086  
ID AAJ04086 standard; peptide; 9 AA.  
XX  
AC AAJ04086;  
XX  
DT 02-JUL-2001 (first entry)  
XX  
DE Hepatitis C virus epitope #4077.  
XX  
HE Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;  
KW antiviral.  
XX  
OS Hepatitis C virus.  
XX  
PN WO200121189-A1.  
XX  
PD 29-MAR-2001.  
XX  
PF 19-JUL-2000; 2000WO-US019774.  
XX  
PR 19-JUL-1999; 99US-00357737.  
XX  
PA (EPIM-) EPIMUNE INC.  
XX  
PI Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;  
FI Baker DM, Celis E, Kubo RT, Grey HM;  
XX  
DR WPI; 2001-308046/32.  
XX  
PT A new composition useful as a vaccines against hepatitis C virus.  
XX  
PS Example 7; Page 203; 214pp; English.

CC The present invention describes a composition comprising a prepared  
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.  
CC These are derived from HCV HLA-binding motifs. They are useful in  
CC vaccines for the prevention and treatment of HCV infection in humans. The  
CC present sequence is an epitope used in the disclosure of the invention  
XX  
SQ Sequence 9 AA;

Query Match 89.4%; Score 42; DB 4; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | |  
Db 1 DLMGYIPLV 9

RESULT 29  
ABG79901  
ID ABG79901 standard; peptide; 9 AA.  
XX  
AC ABG79901;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE MHC class I molecule, viral epitope #149.

XX Major histocompatibility complex; MHC; MHC class I molecule; virus;  
KW epitope; cytotoxic T lymphocyte response; CTL response; lymphatic system;  
KW antigen; immunogenic; malignant tumour; carcinoma; melanoma; leukaemia;  
KW lymphoma; infectious disease; hepatitis; malaria; measles; tuberculosis;  
KW acquired immune deficiency syndrome; AIDS.  
XX  
OS Hepatitis C virus.  
XX  
PN WO200262368-A2.  
XX  
PD 15-AUG-2002.  
XX  
PF 22-JAN-2002; 2002WO-US002033.  
XX  
PR 02-FEB-2001; 2001US-00776232.  
XX  
PA (CTL-) CTL IMMUNOTHERAPIES CORP.  
XX  
PI Kundig TM, Simard J.J.L;  
XX  
PD WPI; 2002-657506/70.  
XX  
PT Inducing or sustaining immunological cytotoxic T lymphocyte response in a  
PT mammal, useful for treating a mammal with malignant tumor or infectious  
PT disease, by directly administering an antigen to the lymphatic system of  
PT the mammal.  
XX  
PS Disclosure; Page 22; 73pp; English.

XX The invention relates to a method of inducing and/or sustaining an  
CC immunological cytotoxic T lymphocyte (CTL) response in a mammal  
CC comprising administering directly to the lymphatic system of the mammal:  
CC (a) an antigen in the form of a polypeptide; (b) a vector comprising a  
CC nucleic acid encoding the antigen; or (c) a non-peptide antigen. The  
CC method is useful for inducing and/or sustaining CTL response in a mammal.  
CC This is particularly useful for treating a mammal having a malignant  
CC tumour (e.g. carcinoma, melanoma, leukaemia or lymphoma) or infectious  
CC disease (e.g. hepatitis, acquired immune deficiency syndrome (AIDS),  
CC malaria, measles or tuberculosis), or in an animal having a  
CC predisposition to these diseases. The mammal may be dogs, cats, mice,  
CC cattle, sheep, pigs, goats, rabbits, or preferably humans. ABG79753-  
CC ABG80319 represent viral epitopes on major histocompatibility complex  
CC (MHC) class I molecules, used in the method of the invention  
XX  
SQ Sequence 9 AA;

Query Match 89.4%; Score 42; DB 5; Length 9;  
Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | | | |  
Db 1 DLMGYIPLV 9

RESULT 30  
ABG80124  
ID ABG80124 standard; peptide; 9 AA.  
XX  
AC ABG80124;  
XX  
DT 29-AUG-2003 (revised)  
DT 15-NOV-2002 (first entry)  
XX  
DE MHC class I molecule, viral epitope #372.  
XX  
KW Major histocompatibility complex; MHC; MHC class I molecule; virus;  
KW epitope; cytotoxic T lymphocyte response; CTL response; lymphatic system;  
KW antigen; immunogenic; malignant tumour; carcinoma; melanoma; leukaemia;  
KW lymphoma; infectious disease; hepatitis; malaria; measles; tuberculosis;  
KW acquired immune deficiency syndrome; AIDS.  
XX

OS Viruses.  
 XX WO200262368-A2.  
 XX  
 XX 15-AUG-2002.  
 XX  
 XX 22-JAN-2002; 2002WO-US002033.  
 XX  
 XX 02-FEB-2001; 2001US-00776232.  
 XX  
 XX (CTL1-) CTL IMMUNOTHERAPIES CORP.  
 XX  
 XX Kundig TM, Simard JLL;  
 XX  
 XX WPI; 2002-657506/70.  
 XX  
 XX Inducing or sustaining immunological cytotoxic T lymphocyte response in a  
 XX mammal, useful for treating a mammal with malignant tumor or infectious  
 XX disease, by directly administering an antigen to the lymphatic system of  
 XX the mammal.  
 XX  
 XX Disclosure; Page 32; 73pp; English.  
 XX  
 XX The invention relates to a method of inducing and/or sustaining an  
 XX immunological cytotoxic T lymphocyte (CTL) response in a mammal  
 XX comprising administering directly to the lymphatic system of the mammal:  
 XX (a) an antigen in the form of a polypeptide; (b) a vector comprising a  
 XX nucleic acid encoding the antigen; or (c) a non-peptide antigen. The  
 XX method is useful for inducing and/or sustaining CTL response in a mammal.  
 XX This is particularly useful for treating a mammal having a malignant  
 XX tumour (e.g. carcinoma, melanoma, leukaemia or lymphoma) or infectious  
 XX disease (e.g. hepatitis, acquired immune deficiency syndrome (AIDS),  
 XX malaria, measles or tuberculosis), or in an animal having a  
 XX predisposition to these diseases. The mammal may be dogs, cats, mice,  
 XX cattle, sheep, pigs, goats, rabbits, or preferably humans. ABG79753-  
 XX ABG80319 represent viral epitopes on major histocompatibility complex  
 XX (MHC) class I molecules, used in the method of the invention. (Updated on  
 XX 29-AUG-2003 to standardise OS field)  
 XX  
 XX Sequence 9 AA;  
 XX

Query Match 89.4%; Score 42; DB 5; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
 |||||  
 Db 1 DLMGYIPLV 9

RESULT 31  
 ABP70576  
 ID ABP70576 standard; peptide; 9 AA.  
 AC ABP70576;

XX 22-APR-2003 (first entry)

XX Hepatitis C virus nucleocapsid protein derived antigenic peptide.

XX Flavivirus; vector; envelope protein; peptide delivery; immune response;  
 XX tumour-associated antigen; immunotherapy; cancer; HCV; nucleocapsid.

XX Hepatitis C virus.

XX WO2002102828-A2.

XX 27-DEC-2002.

XX 31-MAY-2002; 2002WO-US017374.

XX 01-JUN-2001; 2001US-0295265P.

XX

PA (ACAM-) ACAMBIS INC.  
 XX  
 XX Kleanthous H, Oros L, Miller C;  
 XX  
 XX WPI; 2003-167480/16.  
 XX  
 XX Novel flavivirus vector useful for delivering foreign peptides, comprises  
 XX an envelope protein that comprises a foreign peptide.  
 XX  
 XX Example; Page 18; 34pp; English.  
 XX  
 XX The specification describes a flavivirus vector, comprising an envelope  
 XX protein that comprises a foreign peptide. The vector is useful for  
 XX delivery of a peptide, e.g. an antigen to a patient. It can be  
 XX administered to induce an immune response to a pathogen or tumour from  
 XX which the antigen is derived. It also useful for delivering tumour-  
 XX associated antigens for use in immunotherapeutic methods against cancer.  
 XX ABP70571-96 represent B and T cell epitopes from viruses. They may be  
 XX expressed using the vectors of the invention  
 XX  
 XX Sequence 9 AA;  
 XX

Query Match 89.4%; Score 42; DB 6; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
 |||||  
 Db 1 DLMGYIPLV 9

RESULT 32

ADA49528  
 ID ADA49528 standard; peptide; 9 AA.

XX ADA49528;

XX 20-NOV-2003 (first entry)

XX Multi-epitope construct specific epitope #70.

XX multi-epitope; immunogenic; epitope; major histocompatibility complex;  
 XX MHC class I; MHC class II; junctional epitope.

XX Hepatitis C virus.

XX US2002119127-A1.

XX 29-AUG-2002.

XX 27-JUN-2001; 2001US-00894018.

XX 28-DEC-1999; 99US-0173390P.

XX 28-DEC-2000; 2000WO-US035568.

XX 16-APR-2001; 2001US-0284221P.

XX (SETT/) SETTE A.

XX (CHES/) CHESNUT R.

XX (LIVI/) LIVINGSTON B D.

XX (BAKE/) BAKER D M.

XX (NEWM/) NEWMAN M J.

XX (BROW/) BROWN D H.

XX Sette A, Chesnut R, Livingston BD, Baker DM, Newman MJ, Brown DH;  
 XX WPI; 2003-615704/58.

XX Designing multi-epitope construct having major histocompatibility complex  
 XX class I and II epitope nucleic acids, by selecting mixture of amino acid  
 XX insertions at junctions of construct to minimize junctional epitopes.  
 XX  
 XX Disclosure; Fig 19B; 78pp; English.  
 XX



CC The invention relates to a method of designing multi-epitope constructs  
 CC comprising major histocompatibility complex (MHC) class I and II (CTL)  
 CC epitope nucleic acids (CEN), involves sorting CEN, introducing flanking  
 CC amino acid residue selected from specified amino acid residues given in  
 CC specification at C+1 position of CEN, introducing amino acid spacer  
 CC residues between two CEN, and selecting the constructs having less  
 CC junctional epitopes. The method is useful for designing a multi-epitope  
 CC construct having multiple epitope nucleic acid. The method avoids or  
 CC minimises the occurrence of junctional epitopes and maximises the  
 CC immunogenicity and/or antigenicity of multi-epitope vaccines. The present  
 CC sequence represents the amino acid sequence of an epitope present in a  
 CC multi-epitope construct.

XX Sequence 9 AA;

Query Match 89.4%; Score 42; DB 7; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
 DB 1 DLMGYIPLV 9

RESULT 33

AD97587  
 ID ADE97587 standard; peptide; 9 AA.

XX AC ADE97587;

XX DT 12-FEB-2004 (first entry)

XX DE Immunogenic HLA-A2.1 binding peptide #69.

XX cytostatic; anti-inflammatory; hepatotropic; virucide; anti-HIV;  
 XX nephrotropic; neuroprotective; antiarthritic; antirheumatic;  
 XX immunosuppressive; dermatological; muscular; nephrotropic; thyromimetic;  
 XX haemostatic; antithyroid; antianemic; anabolic; hypertensive;  
 XX immunogenic peptide composition; immune response; prostate cancer;  
 XX hepatitis B; hepatitis C; AIDS; renal carcinoma; cervical carcinoma;  
 XX lymphoma; cytomegalovirus; CMV; condyloma acuminatum;  
 XX autoimmunity associated disorder; multiple sclerosis; rheumatoid arthritis;  
 XX Sjogren syndrome; scleroderma; polymyositis; dermatomyositis;  
 XX systemic lupus erythematosus; juvenile rheumatoid arthritis;  
 XX ankylosing spondylitis; myasthenia gravis; MG; bullous pemphigoid;  
 XX pemphigus; glomerulonephritis; Goodpasture's syndrome;  
 XX autoimmune haemolytic anaemia; Hashimoto's disease; pernicious anaemia;  
 XX idiopathic thrombocytopenic purpura; Grave's disease; Addison's disease;  
 XX human leukocyte antigen A2.1; HLA A2.1;  
 XX immunogenic HLA-A2.1 binding peptide.

XX OS Synthetic.

XX PN US2003185822-A1.

XX PD 02-OCT-2003.

XX PF 03-APR-2002; 2002US-00116557.

XX PR 05-MAR-1993; 93US-00027146.

XX PR 04-JUN-1993; 93US-00073205.

XX PR 29-NOV-1993; 93US-00159184.

XX PR 02-DEC-1994; 94US-00349177.

XX PA (GREY/) GREY H M.

XX PA (SETT/) SETTE A.

XX PA (SIDN/) SIDNEY J.

XX PI Grey HM, Sette A, Sidney J;

XX DR WPI; 2004-041186/04.

XX PT Immunogenic peptide composition for preventing, treating or diagnosing

PT pathological states, e.g. prostate cancer, hepatitis B and C, Acquired  
 PT Immunodeficiency Syndrome, and renal carcinoma, includes conserved  
 PT residues at specified positions.

XX Example 11; Page 23; 38pp; English.

XX The invention describes an immunogenic peptide composition comprising 9  
 CC residues including a first conserved residue at a second position from N-  
 CC terminus, and a second conserved residue at C-terminal position. The  
 CC inventive peptide composition is used to elicit an immune response  
 CC against a desired antigen for preventing, treating or diagnosing  
 CC pathological states, e.g. prostate cancer, hepatitis B, hepatitis C,  
 CC AIDS, renal carcinoma, cervical carcinoma, lymphoma, cytomegalovirus  
 CC (CMV), and condyloma acuminatum. It is also used to treat autoimmune  
 CC associated disorders, e.g. multiple sclerosis, rheumatoid arthritis,  
 CC Sjogren syndrome, scleroderma, polymyositis, dermatomyositis, systemic  
 CC lupus erythematosus, juvenile rheumatoid arthritis, ankylosing  
 CC spondylitis, myasthenia gravis (MG), bullous pemphigoid, pemphigus,  
 CC glomerulonephritis, Goodpasture's syndrome, autoimmune hemolytic anemia,  
 CC Hashimoto's disease, pernicious anaemia, idiopathic thrombocytopenic  
 CC purpura, Grave's disease, and Addison's disease. The invention defines  
 CC positions within a motif enabling the selection of the peptides, which  
 CC will bind efficiently to human leukocyte antigen (HLA) A2.1. This is the  
 CC amino acid sequence of an immunogenic HLA-A2.1 binding peptide.

XX Sequence 9 AA;

Query Match 89.4%; Score 42; DB 8; Length 9;

Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
 DB 1 DLMGYIPLV 9

RESULT 34

ADJ79563

ID ADJ79563 standard; peptide; 9 AA.

XX AC ADJ79563;

XX DT 06-MAY-2004 (first entry)

XX CTL epitope for the HCV CORE protein HLA type A2 SeqID 72.

XX bacterial host system; polypeptide; vaccine; cytotoxic T-cell lymphocyte;  
 XX Hepatitis C virus; HCV; infectious mononucleosis;  
 XX nasopharyngeal carcinoma; virucidal; cytostatic; hepatotropic;  
 XX antiinflammatory; anti-HIV; CTL; epitope; NS4.

XX OS Hepatitis C virus.

XX PN WO2004007556-A1.

XX PD 22-JAN-2004.

XX PF 14-JUL-2003; 2003WO-AU000910.

XX PR 12-JUL-2002; 2002AU-00950183.

XX PA (CSLC-) CSL LTD.

XX PA (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX PI Webb EA, Schoofs P;

XX DR WPI; 2004-122896/12.

XX Designing a candidate polypeptide for expression in a host, useful for  
 PT preventing or treating e.g. HIV, comprises identifying hydrophobic  
 PT peptide sequences in the polypeptide and arranging or re-locating the  
 PT peptide sequences.

PS Example 4; SEQ ID NO 72; 98pp; English.

XX This invention relates to a novel method for designing heterologous

CC polypeptides comprising a proportion of hydrophobic amino acids that

CC increases the probability of the protein being efficiently expressed in a

CC bacterial host system. Specifically, it refers to arranging or re-

CC locating a hydrophobic peptide sequence within the protein of interest in

CC order to generate a candidate protein with reduced amplitude in

CC hydrophobicity and/or length of any hydrophobic region. The present

CC invention describes using these polypeptides for use in a polypeptide

CC vaccine that is capable of eliciting a cytotoxic T-cell lymphocyte (CTL)

CC immune response. As such, compositions can be used for the prevention or

CC treatment of diseases associated with the Epstein-Barr virus (EBV) such

CC as infectious mononucleosis or nasopharyngeal carcinoma, as well as

CC hepatitis C virus, cytomegalovirus or HIV. Accordingly, they exhibit

CC virucidal, cytostatic, hepatotropic, antiinflammatory and anti-HIV

CC activities. This peptide sequence is a CTL epitope that can be included

CC in the HCV vaccine of the invention.

XX

SQ Sequence 9 AA;

Query Match 89.4%; Score 42; DB 8; Length 9;

Best Local Similarity 88.9%; Pred. No. 1.8e+06; Mismatches 1; Indels 0; Gaps 0;

Matches 8; Conservative 0;

Qy 1 DLMGYIPAV 9

Db 1 DLMGYIPLV 9

|||||||

RESULT 35

ADK69063

ID ADK69063 standard; peptide; 9 AA.

XX

AC ADK69063;

XX

DT 06-MAY-2004 (first entry)

XX

DE Epitope liberation-related peptide SeqID426.

XX

KW epitope liberation; substrate; proteasome; cytostatic; antibacterial;

KW protozoicide; fungicide; T-cell activator; vaccine; housekeeping epitope;

KW cytotoxic T lymphocyte; CTL; adoptive immunotherapy; neoplastic cell;

KW virus; bacterium; protozoan; fungus; housekeeping proteasome system.

XX

OS Human herpesvirus 5.

XX

XX

PN US2003228634-A1.

XX

PD 11-DEC-2003.

XX

PF 07-NOV-2002; 2002US-00292413.

XX

PR 07-NOV-2001; 2001US-0336968P.

XX

PA (SIMA/) SIMARD J J L.

PA (DIAM/) DIAMOND D C.

PA (QIUZ/) QIU Z.

PA (LEIX/) LEI X.

XX

PI Simard J J L, Diamond DC, Qiu Z, Lei X;

XX

DR WPI; 2004-167209/16.

XX

XX Identifying polypeptide suitable for epitope e.g., housekeeping epitope,

PT liberation by contacting substrate polypeptide comprising epitope of

PT interest, with proteasome, and assaying for liberation of epitope.

XX

PS Disclosure; SEQ ID NO 426; 67pp; English.

XX

XX This invention relates to a novel method of identifying a polypeptide

CC suitable for epitope liberation, including the steps of identifying an

CC epitope of interest, providing substrate polypeptide sequence including

CC

the epitope, wherein the substrate permits processing by a proteasome;

contacting the substrate with a composition including the proteasome,

under conditions that support processing of the substrate by proteasome,

and assaying for liberation of epitope. The invention may be useful for

the development of compounds with a cytostatic, antibacterial,

protozoicide or fungicide activity acting as T-cell activators. In

addition, the invention may allow development of a vaccine. The invention

is useful for identifying a polypeptide suitable for epitope liberation,

where the epitope is a housekeeping epitope. The compositions comprising

the identified housekeeping epitopes are useful in vitro in vaccine

development or in the generation or expansion of cytotoxic T lymphocyte

(CTL) to be used in adoptive immunotherapy. The invention is also useful

for activating T-cells against neoplastic cells, and cells infected with

virus, bacterium, protozoan or fungus. CTL epitopes are identified based

on the knowledge that such epitopes are, in fact, produced by the

housekeeping proteasome system. Once identified, these epitopes, embodied

as peptides, can be used to successfully immunise or induce therapeutic

CTL responses against housekeeping proteasome expressing target cells in

the host. The present sequence is that of a peptide which is related to

the method of the invention.

XX

SQ Sequence 9 AA;

Query Match 89.4%; Score 42; DB 8; Length 9;

Best Local Similarity 88.9%; Pred. No. 1.8e+06; Mismatches 1; Indels 0; Gaps 0;

Matches 8; Conservative 0;

Qy 1 DLMGYIPAV 9

Db 1 DLMGYIPLV 9

|||||||

RESULT 36

ADK68837

ID ADK68837 standard; peptide; 9 AA.

XX

AC ADK68837;

XX

DT 06-MAY-2004 (first entry)

XX

DE Epitope liberation-related peptide SeqID200.

XX

KW epitope liberation; substrate; proteasome; cytostatic; antibacterial;

KW protozoicide; fungicide; T-cell activator; vaccine; housekeeping epitope;

KW cytotoxic T lymphocyte; CTL; adoptive immunotherapy; neoplastic cell;

KW virus; bacterium; protozoan; fungus; housekeeping proteasome system.

XX

OS Hepatitis C virus.

XX

XX

PN US2003228634-A1.

XX

PD 11-DEC-2003.

XX

PF 07-NOV-2002; 2002US-00292413.

XX

PR 07-NOV-2001; 2001US-0336968P.

XX

PA (SIMA/) SIMARD J J L.

PA (DIAM/) DIAMOND D C.

PA (QIUZ/) QIU Z.

PA (LEIX/) LEI X.

XX

PI Simard J J L, Diamond DC, Qiu Z, Lei X;

XX

DR WPI; 2004-167209/16.

XX

XX Identifying polypeptide suitable for epitope e.g., housekeeping epitope,

PT liberation by contacting substrate polypeptide comprising epitope of

PT interest, with proteasome, and assaying for liberation of epitope.

XX

PS Disclosure; SEQ ID NO 200; 67pp; English.

XX

XX This invention relates to a novel method of identifying a polypeptide

CC suitable for epitope liberation, including the steps of identifying an  
 CC epitope of interest; providing substrate polypeptide sequence including  
 CC the epitope, wherein the substrate permits processing by a proteasome;  
 CC contacting the substrate with a composition including the proteasome;  
 CC under conditions that support processing of the substrate by proteasome;  
 CC and assaying for liberation of epitope. The invention may be useful for  
 CC the development of compounds with a cytostatic, antibacterial,  
 CC protozoacide or fungicide activity acting as T-cell activators. In  
 CC addition, the invention may allow development of a vaccine. The invention  
 CC is useful for identifying a polypeptide suitable for epitope liberation,  
 CC where the epitope is a housekeeping epitope. The compositions comprising  
 CC the identified housekeeping epitopes are useful in vitro in vaccine  
 CC development or in the generation or expansion of cytotoxic T lymphocyte  
 CC (CTL) to be used in adoptive immunotherapy. The invention is also useful  
 CC for activating T-cells against neoplastic cells, and cells infected with  
 CC virus, bacterium, protozoan or fungus. CTL epitopes are identified based  
 CC on the knowledge that such epitopes are, in fact, produced by the  
 CC housekeeping proteasome system. Once identified, these epitopes, embodied  
 CC as peptides, can be used to successfully immunise or induce therapeutic  
 CC CTL responses against housekeeping proteasome expressing target cells in  
 CC the host. The present sequence is that of a peptide which is related to  
 CC the method of the invention.

SQ Sequence 9 AA;

Query Match 89.4%; Score 42; DB 8; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
 |||||  
 Db 1 DLMGYIPLV 9

# RESULT 37

ID ADJ84718 standard; peptide; 9 AA.

AC ADJ84718;

DT 06-MAY-2004 (first entry)

XX Hepatitis C virus core protein CTL epitope SEQ ID NO:176.

XX lipopeptide; lipid moiety; T helper cell epitope; Th epitope;  
 KW cytotoxic T cell epitope; CTL epitope; immune response; immunisation;  
 KW influenza; hepatitis C virus; Listeria monocytogenes; vaccine; cancer;  
 KW cytostatic; virucide.

OS Hepatitis C virus.  
 OS Synthetic.

PN WO2004014957-A1.

PD 19-FEB-2004.

PF 12-AUG-2003; 2003WO-AU001019.

PR 12-AUG-2002; 2002US-0403328P.

XX (COUN-) COUNCIL QUEENSLAND INST MEDICAL RES.

XX Jackson D, Zeng W;

XX WPI; 2004-238736/22.

XX New lipopeptide comprising a polypeptide comprising an amino acid  
 PT sequence of a T helper cell and cytotoxic T cell epitope, useful for  
 PT preparing a composition for treating or preventing cancer, or hepatitis C  
 PT virus or influenza virus.

PS Claim 22; SEQ ID NO 176; 166pp; English.

CC The present invention describes a lipopeptide comprising a polypeptide  
 CC conjugated to one or more lipid moieties, where the polypeptide comprises  
 CC an amino acid sequence that comprises: (a) the amino acid sequence of a T  
 CC helper cell (Th) epitope and the amino acid sequence of a cytotoxic T  
 CC cell (CTL) epitope, where the amino acid sequences are different; and (b)  
 CC one or more internal lysine residues or internal lysine analogues  
 CC for covalent attachment of each of the lipid moieties via the epsilon-  
 CC amino group or terminal side-chain group of the lysine or lysine  
 CC analogue. Each of the one or more lipid moieties is covalently attached  
 CC to an epsilon-amino group of the one or more internal lysine residues or  
 CC to a terminal side-chain group of the one or more internal lysine  
 CC analogue residues. Also described: (1) producing a lipopeptide; (2) a  
 CC composition comprising the lipopeptide and an excipient or diluent; (3)  
 CC eliciting an immune response in a subject; (4) immunising a subject  
 CC against influenza or hepatitis C virus or Listeria monocytogenes; (5) a  
 CC vaccine against an influenza virus or hepatitis C virus or Listeria  
 CC monocytogenes comprising the lipopeptides, where the CTL epitope is from  
 CC an influenza virus or hepatitis C virus or Listeria monocytogenes protein  
 CC ; (6) treating or preventing cancer; and (7) a vaccine against cancer  
 CC comprising the lipopeptide, where the CTL epitope is a tumour-specific  
 CC CTL epitope. The lipopeptide has cytostatic and virucide activities, and  
 CC can be used in vaccine. The lipopeptide is useful for preparing a  
 CC composition for treating or preventing cancer, or hepatitis C virus or  
 CC influenza virus. The present sequence is used in the exemplification of  
 CC the present invention.

SQ Sequence 9 AA;

Query Match 89.4%; Score 42; DB 8; Length 9;  
 Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9  
 |||||  
 Db 1 DLMGYIPLV 9

# RESULT 38

ID ADO24206 standard; peptide; 9 AA.

AC ADO24206;

XX 01-JUL-2004 (first entry)

DT HCV epitope peptide #5 for multi-epitope construct.

XX hepatotropic; virucide; antiinflammatory; anti-HIV; cytostatic; epitope;  
 KW cytotoxic T lymphocyte; CTL; helper T lymphocyte; HTL; immune response;  
 KW hepatitis B virus; HBV; HIV; cancer; lymphoma.

XX Hepatitis C virus.

XX WO2004031210-A2.

XX 15-APR-2004.

XX 03-OCT-2003; 2003WO-US031303.

XX 03-OCT-2002; 2002US-0415463P.

XX 22-OCT-2002; 2002US-0419973P.

XX (EPIM-) EPIMMUNE INC.

XX (GEMV) GENENCOR INT. INC.

XX Sette A, Chesnut R, Newman MJ, Livingston BD, Babe LM, Chen Y;  
 XX Deyoung LM, Huang MTF, Power SD;

XX WPI; 2004-330143/30.

XX New multi-epitope polynucleotides encoding cytotoxic T lymphocyte and/or  
 PT helper T lymphocyte epitopes, useful for inducing or stimulating an  
 PT antiviral or anticancer immune response, especially against hepatitis B

```

PT virus.
XX Disclosure, SEQ ID NO 356; 401pp; English.
XX
XX
CC The invention relates to a multi-epitope polynucleotide (I) encoding
CC cytotoxic T lymphocyte (CTL) and/or helper T lymphocyte (HTL) epitopes.
CC The composition and methods are useful for inducing or stimulating an
CC immune response against hepatitis B virus (HBV), other viruses (e.g.
CC HIV), or against cancers such as prostate cancer or lymphoma. This
CC sequence corresponds to an example of an epitope peptide used for multi-
CC epitope construct.
XX
XX
SQ Sequence 9 AA;
Query Match 89.4%; Score 42; DB 8; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.8e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9
Db 1 DLMGYIPLV 9

RESULT 39
AD008619
ID AD008619 standard; peptide; 9 AA.
XX
XX AD008619;
XX
XX
DT 12-AUG-2004 (first entry)
DE Hepatitis C virus core epitope peptide SEQ ID NO:2.
XX
XX immunogenic; immunodominant; Hepatitis C Virus; HCV;
XX human leukocyte antigen class I molecule; HLA class I molecule;
XX HCV-specific immune response; immune response;
XX immunodominant HCV epitope; virucide; hepatotropic; vaccine;
XX HCV infection; infection; epitope.
XX
XX Hepatitis C virus.
XX WO2004041842-A2.
XX
XX 21-MAY-2004.
XX
XX 16-MAY-2003; 2003WO-US015443.
XX
XX 16-MAY-2002; 2002US-0381273P.
XX
XX (GEO ) GEN HOSPITAL CORP.
XX
XX Lauer G, Ouchi K, Walker BD;
XX WPI; 2004-400639/37.
XX
XX New immunogenic composition comprising immunodominant Hepatitis C Virus
XX (HCV) polypeptides that bind to a spectrum of HLA class I molecules,
XX useful in treating or preventing HCV infection.
XX
XX Disclosure; SEQ ID NO 2; 68pp; English.
XX
XX The present invention describes an immunogenic composition (I) comprising
XX immunodominant Hepatitis C Virus (HCV) polypeptides that bind to a
XX spectrum of human leukocyte antigen (HLA) class I molecules that are
XX expressed in at least 85 % of the individuals in a target population.
XX Also described: (1) stimulating an HCV-specific immune response; (2)
XX measuring an immune response in a patient; and (3) identifying an
XX immunodominant HCV epitope. (I) has virucide and hepatotropic activities,
XX and can be used in vaccines. The immunogenic composition (I) is useful in
XX treating or preventing HCV infection. The present sequence represents an
XX HCV epitope peptide, which is used in the exemplification of the present
XX invention.
XX
XX
SQ Sequence 9 AA;
Query Match 89.4%; Score 42; DB 8; Length 9;
Best Local Similarity 88.9%; Pred. No. 1.8e+06;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DLMGYIPAV 9
Db 1 DLMGYIPLV 9

RESULT 40
ADQ10635
ID ADQ10635 standard; peptide; 9 AA.
XX
XX ADQ10635;
XX
XX 23-SEP-2004 (first entry)
DE Hepatitis C virus T-cell epitope seqid 200.
XX
XX immunostimulant; cytostatic; vaccine; tumour-associated antigen SSX-2;
XX SSX-2 antigen; epitope cluster; MHC receptor peptide binding cleft;
XX immunogenic composition; immune response; cancer; vaccine vector;
XX epitope liberation; human leukocyte antigen; HLA A2-specific CTL;
XX cytotoxic T lymphocyte; T-cell epitope.
XX
XX Hepatitis C virus.
XX US2004132088-A1.
XX
XX 08-JUL-2004.
XX
XX 10-FEB-2004; 2004US-00777053.
XX
XX 07-NOV-2001; 2001US-0336968P.
XX
XX 07-NOV-2002; 2002US-00292413.
XX
XX (SIMA/) SIMARD J J L.
XX (DIAM/) DIAMOND D C.
XX (QIUZ/) QIU Z.
XX (LEIX/) LEI X.
XX
XX Simard J J L, Diamond DC, Qiu Z, Lei X;
XX WPI; 2004-517003/49.
XX
XX Novel nucleic acid encoding tumor-associated antigen SSX-2, useful in
XX inducing an immune response and in treating cancer.
XX
XX Disclosure; SEQ ID NO 200; 260pp; English.
XX
XX The invention describes an isolated nucleic acid (I) comprising a reading
XX frame comprising a first sequence, where the first sequence encodes one
XX or more segments of tumour-associated antigen SSX-2, which comprises a
XX sequence of 188 amino acids (SEQ ID NO: 40), where the first sequence
XX does not encode the complete SSX-2 antigen, and where each segment
XX comprises an epitope cluster, the cluster comprising or encoding at least
XX two amino acid sequences having a known or predicted affinity for a same
XX MHC receptor peptide binding cleft. Also described are: an isolated
XX polypeptide comprising the amino acid sequence encoded in the reading
XX frame; and an immunogenic composition comprising (I) or the polypeptide
XX of (I). (I) is a nucleic acid encoding a tumour-associated antigen SSX-2
XX comprising a fully defined sequence of 188 amino acids (SEQ ID NO: 40).
XX The nucleic acid, the encoded antigen, and composition are useful in
XX inducing an immune response and in treating cancer. Expression cassettes
XX are used in vaccine vectors. This is the amino acid sequence of a T-cell
XX epitope MHC ligand associated with methods, therapies and compositions
XX described in the invention.
XX
XX
SQ Sequence 9 AA;
Query Match 89.4%; Score 42; DB 8; Length 9;

```

Best Local Similarity 88.9%; Pred. No. 1.8e+06;  
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DLMGYIPAV 9  
| | | | |  
Db 1 DLMGYIPLV 9

Search completed: August 23, 2005, 13:32:11  
Job time : 164 secs

***This Page Blank (usptq)***